

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
21 June 2001 (21.06.2001)

PCT

(10) International Publication Number
WO 01/44448 A2

(51) International Patent Classification⁷: C12N 9/02,
15/11, A61K 38/44

Valley, CA 94552 (US). LU, Dyung, Aina, M. [US/US];
233 Coy Drive, San Jose, CA 95123 (US).

(21) International Application Number: PCT/US00/33158

(74) Agents: HAMLET-COX, Diana et al.; Incyte Genomics,
Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).

(22) International Filing Date: 7 December 2000 (07.12.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/172,367 16 December 1999 (16.12.1999) US

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ,
DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR,
HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR,
LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ,
NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM,
TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(71) Applicant (*for all designated States except US*): INCYTE
GENOMICS, INC. [US/US]; 3160 Porter Drive, Palo
Alto, CA 94304 (US).

(84) Designated States (*regional*): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,
IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,
CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): YUE, Henry
[US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US).
LAL, Preeti [IN/US]; 2382 Lass Drive, Santa Clara, CA
95054 (US). TANG, Y., Tom [CN/US]; 4230 Ranwick
Court, San Jose, CA 95118 (US). HILLMAN, Jennifer,
L. [US/US]; 230 Monroe Drive #17, Mountain View, CA
94040 (US). BAUGHN, Mariah, R. [US/US]; 14244
Santiago Road, San Leandro, CA 94577 (US). AZIMZAI,
Yalda [US/US]; 5518 Boulder Canyon Drive, Castro

Published:

— Without international search report and to be republished
upon receipt of that report.

For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.

(54) Title: HUMAN OXIDOREDUCTASE PROTEINS

(57) Abstract: The invention provides human oxidoreductase proteins (ORP) and polynucleotides which identify and encode ORP. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with expression of ORP.

WO 01/44448 A2

HUMAN OXIDOREDUCTASE PROTEINS

TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of oxidoreductase proteins and to the use of these sequences in the diagnosis, treatment, and prevention of cell proliferative disorders including cancer; endocrine, metabolic, reproductive, neurological, viral, and autoimmune/inflammatory disorders, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of oxidoreductase proteins.

BACKGROUND OF THE INVENTION

Many pathways of biogenesis and biodegradation require oxidoreductase (dehydrogenase or reductase) activity, coupled to the reduction or oxidation of a donor or acceptor cofactor. Potential cofactors include cytochromes, oxygen, disulfide, iron-sulfur proteins, flavin adenine dinucleotide (FAD), and the nicotinamide adenine dinucleotides NAD and NADP (Newsholme, E.A. and Leech, A.R. (1983) Biochemistry for the Medical Sciences, John Wiley and Sons, Chichester, U.K. pp. 779-793).

Reductase activity catalyzes the transfer of electrons between substrate(s) and cofactor(s) with concurrent oxidation of the cofactor. The reverse dehydrogenase reaction catalyzes the reduction of a cofactor and consequent oxidation of the substrate. Oxidoreductase enzymes are a broad superfamily of proteins that catalyze numerous reactions in all cells of organisms ranging from bacteria to plants to humans. These reactions include metabolism of sugar, certain detoxification reactions in the liver, and the synthesis or degradation of fatty acids, amino acids, glucocorticoids, estrogens, androgens, and prostaglandins. Different family members are named according to the direction in which their reactions are typically catalyzed; thus they may be referred to as oxidoreductases, oxidases, reductases, or dehydrogenases. In addition, family members often have distinct cellular localizations, including the cytosol, the plasma membrane, mitochondrial inner or outer membrane, and peroxisomes.

Tetrahydrofolate is a derivatized glutamate molecule that acts as a carrier, providing activated one-carbon units to a wide variety of biosynthetic reactions, including synthesis of purines, pyrimidines, and the amino acid methionine. Tetrahydrofolate is generated by the activity of a holoenzyme complex called tetrahydrofolate synthase, which includes three enzyme activities: tetrahydrofolate dehydrogenase, tetrahydrofolate cyclohydrolase, and tetrahydrofolate synthetase. Thus, tetrahydrofolate dehydrogenase plays an important role in generating building blocks for nucleic and amino acids, crucial to proliferating cells.

Intracellular redox status plays a critical role in the assembly of proteins. A major rate limiting step in protein folding is the thiol:disulfide exchange necessary for correct protein assembly. Although

incubation of reduced, unfolded proteins in buffers containing defined ratios of oxidized and reduced thiols can lead to folding into native conformation, the rate of folding is slow, and the attainment of the native conformation decreases proportionately with protein size and the number of cysteine residues. Certain cellular compartments such as the endoplasmic reticulum of eukaryotes and the periplasmic space of prokaryotes are maintained in a more oxidized state than the surrounding cytosol. Correct disulfide formation can occur in these compartments, but it occurs at a rate that is insufficient for normal cell processes and inadequate for synthesizing secreted proteins.

Protein disulfide isomerases (PDIs), thioredoxins, and glutaredoxins are able to catalyze the formation of disulfide bonds and regulate the redox environment in cells to enable the necessary thiol:disulfide exchanges. Each of these classes of molecules has a somewhat different function, but all belong to a group of disulfide-containing redox proteins that contain a conserved active-site sequence and are ubiquitously distributed in eukaryotes and prokaryotes. PDIs are found in the endoplasmic reticulum of eukaryotes and in the periplasmic space of prokaryotes. PDIs function by exchanging their own disulfide for thiols in a folding peptide chain. In contrast, reduced thioredoxins and glutaredoxins are generally found in the cytoplasm and function by directly reducing disulfides in the substrate proteins. Thioredoxin (Trx), a heat-stable, redox-active protein, contains an active site cysteine disulfide/dithiol. Oxidized thioredoxin, Trx-S, can be reduced to the dithiol form by NADPH and a specific flavoprotein enzyme, thioredoxin reductase. Reduced thioredoxin, Trx-(SH), participates in a number of redox reactions mostly linked to reduction of protein disulfides. Trx and thioredoxin reductase (TR), together with NADPH, form a redox complex in which TR catalyzes the electron transport from NADPH to Trx. The reduced thioredoxin then functions as an electron donor in a wide variety of different metabolic processes.

Disulfide-containing redox proteins not only facilitate disulfide formation, but also regulate and participate in a wide variety of physiological processes. The thioredoxin system serves, for example, as a hydrogen donor for ribonucleotide reductase and controls the activity of enzymes by redox reactions. Mammalian thioredoxin (MT) acts as a hydrogen donor for ribonucleotide reductase and methionine sulfoxide reductase, facilitates refolding of disulfide-containing proteins, and activates the glucocorticoid and interleukin-2 receptors. MT also modulates the DNA binding activity of some transcription factors either directly (TFIIIC, BZLF1, and NF- κ B) or indirectly (AP-1) through the nuclear factor Ref-1. The importance of the redox regulation of transcription factors is exemplified by the v-fos oncogene where a point mutation of the thioredoxin-modulated cysteine residue results in constitutive activation of the AP-1 complex. Thioredoxin, secreted by cells using a leaderless pathway, stimulates the proliferation of lymphoid cells, fibroblasts, and a variety of human solid tumor cell lines. Furthermore, thioredoxin is an essential component of early pregnancy factor, inhibits human immunodeficiency virus expression in macrophages, reduces H_2O_2 , scavenges free radicals, and protects

cells against oxidative stress (Abate, C. et al. (1990) *Science* 249: 1157-1161; Rosen, A. et al. (1995) *Int. Immunol.* 7: 625-633; Tagaya, Y. et al. (1989) *EMBO J.* 8: 757-764; Newman, G. W. (1994) *J. Expt. Med.* 180: 359-363; and Makino, Y. (1996) *J. Clin. Invest.* 98: 2469-2477).

Short-chain alcohol dehydrogenases (SCADs) are a family of dehydrogenases that share only 15% to 30% sequence identity, with similarity predominantly in the coenzyme binding domain and the substrate binding domain. In addition to the well-known role in detoxification of ethanol, SCADs are also involved in synthesis and degradation of fatty acids, steroids, and some prostaglandins, and are therefore implicated in a variety of disorders such as lipid storage disease, myopathy, SCAD deficiency, and certain genetic disorders. For example, retinol dehydrogenase is a SCAD-family member (Simon, A. et al. (1995) *J. Biol. Chem.* 270:1107-1112) that converts retinol to retinal, the precursor of retinoic acid. Retinoic acid, a regulator of differentiation and apoptosis, has been shown to down-regulate genes involved in cell proliferation and inflammation (Chai, X. et al. (1995) *J. Biol. Chem.* 270:3900-3904). In addition, retinol dehydrogenase has been linked to hereditary eye diseases such as autosomal recessive childhood-onset severe retinal dystrophy (Simon, A. et al. (1996) *Genomics* 36:424-430).

Propagation of nerve impulses, modulation of cell proliferation and differentiation, induction of the immune response, and tissue homeostasis involve neurotransmitter metabolism (Weiss, B. (1991) *Neurotoxicology* 12:379-386; Collins, S.M. et al. (1992) *Ann. N.Y. Acad. Sci.* 664:415-424; Brown, J.K. and Imam, H. (1991) *J. Inherit. Metab. Dis.* 14:436-458). Many pathways of neurotransmitter metabolism require oxidoreductase activity, coupled to reduction or oxidation of a cofactor, such as NAD⁺/NADH (Newsholme, E.A. and Leech, A.R. (1983) *Biochemistry for the Medical Sciences*, John Wiley and Sons, Chichester, U.K. pp. 779-793). Degradation of catecholamines (epinephrine or norepinephrine) requires alcohol dehydrogenase (in the brain) or aldehyde dehydrogenase (in peripheral tissue). NAD⁺-dependent aldehyde dehydrogenase oxidizes 5-hydroxyindole-3-acetate (the product of 5-hydroxytryptamine (serotonin) metabolism) in the brain, blood platelets, liver and pulmonary endothelium (Newsholme, E.A. and Leech, A.R. *supra* p. 786). Other neurotransmitter degradation pathways that utilize NAD⁺/NADH-dependent oxidoreductase activity include those of L-DOPA (precursor of dopamine, a neuronal excitatory compound), glycine (an inhibitory neurotransmitter in the brain and spinal cord), histamine (liberated from mast cells during the inflammatory response), and taurine (an inhibitory neurotransmitter of the brain stem, spinal cord and retina) (Newsholme, E.A. and Leech, A.R. *supra*, pp. 790, 792). Epigenetic or genetic defects in neurotransmitter metabolic pathways can result in a spectrum of disease states in different tissues including Parkinson's disease and inherited myoclonus (McCance, K.L. and Huether, S.E. (1994) *Pathophysiology*, Mosby-Year Book, Inc., St. Louis, MO pp. 402-404; Gundlach, A.L. (1990) *FASEB J.* 4:2761-2766).

3-Hydroxyacyl-CoA dehydrogenase (3HACD) is involved in fatty acid metabolism. It catalyzes the reduction of 3-hydroxyacyl-CoA to 3-oxoacyl-CoA, with concomitant oxidation of NAD

to NADH, in the mitochondria and peroxisomes of eukaryotic cells. In peroxisomes, 3HACD and enoyl-CoA hydratase form an enzyme complex called bifunctional enzyme, defects in which are associated with peroxisomal bifunctional enzyme deficiency. This interruption in fatty acid metabolism produces accumulation of very-long chain fatty acids, disrupting development of the brain, bone, and adrenal glands. Infants born with this deficiency typically die within 6 months (Watkins, P. et al. (1989) J. Clin. Invest. 83:771-777; Online Mendelian Inheritance in Man (OMIM), #261515). The neurodegeneration that is characteristic of Alzheimer's disease involves development of extracellular plaques in certain brain regions. A major protein component of these plaques is the peptide amyloid- β (A β), which is one of several cleavage products of amyloid precursor protein (APP). 3HACD, which has been shown to bind the A β peptide, is overexpressed in neurons affected in Alzheimer's disease. In addition, an antibody against 3HACD can block the toxic effects of A β in a cell culture model of Alzheimer's disease (Yan, S. et al. (1997) Nature 389:689-695; OMIM, #602057).

17 β -hydroxysteroid dehydrogenase (17 β HSD6) plays an important role in the regulation of the male reproductive hormone, dihydrotestosterone (DHTT). 17 β HSD6 acts to reduce levels of DHTT by oxidizing a precursor of DHTT, 3 α -diol, to androsterone which is readily glucuronidated and removed from tissues. 17 β HSD6 is active with both androgen and estrogen substrates when expressed in embryonic kidney 293 cells. At least five other isozymes of 17 β HSD have been identified that catalyze oxidation and/or reduction reactions in various tissues with preferences for different steroid substrates (Biswas, M.G. and Russell, D.W. (1997) J. Biol. Chem. 272:15959-15966). For example, 17 β HSD1 preferentially reduces estradiol and is abundant in the ovary and placenta. 17 β HSD2 catalyzes oxidation of androgens and is present in the endometrium and placenta. 17 β HSD3 is exclusively a reductive enzyme in the testis (Geissler, W.M. et al. (1994) Nature Genet. 7:34-39). An excess of androgens such as DHTT can contribute to certain disease states such as benign prostatic hyperplasia and prostate cancer.

Steroids, such as estrogen, testosterone, corticosterone, and others, are generated from a common precursor, cholesterol, and are interconverted into one another. A wide variety of enzymes act upon cholesterol, including a number of dehydrogenases. One such dehydrogenase is 3-oxo-5- α -steroid dehydrogenase (OASD), a microsomal membrane protein highly expressed in prostate and other androgen-responsive tissues. OASD catalyzes the conversion of testosterone into dihydrotestosterone, which is the most potent androgen. Dihydrotestosterone is essential for the formation of the male phenotype during embryogenesis, as well as for proper androgen-mediated growth of tissues such as the prostate and male genitalia. A defect in OASD that prevents the conversion of testosterone into dihydrotestosterone leads to a rare form of male pseudohermaphroditis, characterized by defective formation of the external genitalia (Andersson, S., et al. (1991) Nature 354:159-161; Labrie, F., et al. (1992) Endocrinology 131:1571-1573; OMIM #264600). Thus, OASD plays a central role in sexual

differentiation and androgen physiology.

The discovery of new oxidoreductase proteins and the polynucleotides encoding them satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of cell proliferative disorders including cancer; endocrine, metabolic, reproductive, neurological, viral, and autoimmune/inflammatory disorders, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of oxidoreductase proteins.

SUMMARY OF THE INVENTION

The invention features purified polypeptides, oxidoreductase proteins, referred to collectively as "ORP" and individually as "ORP-1," "ORP-2," "ORP-3," "ORP-4," "ORP-5," "ORP-6," "ORP-7," "ORP-8," "ORP-9," "ORP-10," "ORP-11," "ORP-12," "ORP-13," "ORP-14," "ORP-15," "ORP-16," "ORP-17," "ORP-18," "ORP-19," "ORP-20," "ORP-21," "ORP-22," "ORP-23," "ORP-24," "ORP-25," "ORP-26," "ORP-27." In one aspect, the invention provides an isolated polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-27. In one alternative, the invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1-27.

The invention further provides an isolated polynucleotide encoding a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-27. In one alternative, the polynucleotide encodes a polypeptide selected from the group consisting of SEQ ID NO:1-27. In another alternative, the polynucleotide is selected from the group consisting of SEQ ID NO:28-54.

Additionally, the invention provides a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, c)

a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-27. In one alternative, the invention provides a cell transformed with the recombinant polynucleotide. In another alternative, the invention provides a transgenic organism comprising the recombinant polynucleotide.

The invention also provides a method for producing a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-27. The method comprises a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding the polypeptide, and b) recovering the polypeptide so expressed.

Additionally, the invention provides an isolated antibody which specifically binds to a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-27.

The invention further provides an isolated polynucleotide comprising a polynucleotide sequence selected from the group consisting of a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:28-54, b) a naturally occurring polynucleotide sequence having at least 70% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:28-54, c) a polynucleotide sequence complementary to a), d) a polynucleotide sequence complementary to b), and e) an RNA equivalent of a)-d). In one alternative, the polynucleotide comprises at least 60 contiguous nucleotides.

Additionally, the invention provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide comprising a polynucleotide sequence selected from the group consisting of a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:28-54, b) a naturally occurring polynucleotide sequence having at least 70% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:28-54, c) a polynucleotide sequence complementary to a), d) a polynucleotide sequence complementary to b), and e)

an RNA equivalent of a)-d). The method comprises a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and b) detecting the presence or absence of said hybridization complex, and optionally, if present, the amount thereof. In one alternative, the probe comprises at least 60 contiguous nucleotides.

The invention further provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide comprising a polynucleotide sequence selected from the group consisting of a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:28-54, b) a naturally occurring polynucleotide sequence having at least 70% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:28-54, c) a polynucleotide sequence complementary to a), d) a polynucleotide sequence complementary to b), and e) an RNA equivalent of a)-d). The method comprises a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

The invention further provides a composition comprising an effective amount of a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, and a pharmaceutically acceptable excipient. In one embodiment, the composition comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-27. The invention additionally provides a method of treating a disease or condition associated with decreased expression of functional ORP, comprising administering to a patient in need of such treatment the composition.

The invention also provides a method for screening a compound for effectiveness as an agonist of a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-27. The method comprises a)

exposing a sample comprising the polypeptide to a compound, and b) detecting agonist activity in the sample. In one alternative, the invention provides a composition comprising an agonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with decreased expression of functional ORP, comprising administering to a patient in need of such treatment the composition.

Additionally, the invention provides a method for screening a compound for effectiveness as an antagonist of a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-27. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting antagonist activity in the sample. In one alternative, the invention provides a composition comprising an antagonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with overexpression of functional ORP, comprising administering to a patient in need of such treatment the composition.

The invention further provides a method of screening for a compound that specifically binds to a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-27. The method comprises a) combining the polypeptide with at least one test compound under suitable conditions, and b) detecting binding of the polypeptide to the test compound, thereby identifying a compound that specifically binds to the polypeptide.

The invention further provides a method of screening for a compound that modulates the activity of a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-27. The

method comprises a) combining the polypeptide with at least one test compound under conditions permissive for the activity of the polypeptide, b) assessing the activity of the polypeptide in the presence of the test compound, and c) comparing the activity of the polypeptide in the presence of the test compound with the activity of the polypeptide in the absence of the test compound, wherein a
5 change in the activity of the polypeptide in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide.

The invention further provides a method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence selected from the group consisting of SEQ ID NO:28-54, the method comprising a)
10 exposing a sample comprising the target polynucleotide to a compound, and b) detecting altered expression of the target polynucleotide.

The invention further provides a method for assessing toxicity of a test compound, said method comprising a) treating a biological sample containing nucleic acids with the test compound; b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20
15 contiguous nucleotides of a polynucleotide comprising a polynucleotide sequence selected from the group consisting of i) a polynucleotide sequence selected from the group consisting of SEQ ID NO:28-54, ii) a naturally occurring polynucleotide sequence having at least 70% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:28-54, iii) a polynucleotide sequence complementary to i), iv) a polynucleotide sequence complementary to ii),
20 and v) an RNA equivalent of i)-iv). Hybridization occurs under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide comprising a polynucleotide sequence selected from the group consisting of i) a polynucleotide sequence selected from the group consisting of SEQ ID NO:28-54, ii) a naturally occurring polynucleotide sequence having at least 70% sequence identity to a
25 polynucleotide sequence selected from the group consisting of SEQ ID NO:28-54, iii) a polynucleotide sequence complementary to i), iv) a polynucleotide sequence complementary to ii), and v) an RNA equivalent of i)-iv). Alternatively, the target polynucleotide comprises a fragment of a polynucleotide sequence selected from the group consisting of i)-v) above; c) quantifying the amount of hybridization complex; and d) comparing the amount of hybridization complex in the
30 treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

BRIEF DESCRIPTION OF THE TABLES

35 Table 1 shows polypeptide and nucleotide sequence identification numbers (SEQ ID NOs),

clone identification numbers (clone IDs), cDNA libraries, and cDNA fragments used to assemble full-length sequences encoding ORP.

Table 2 shows features of each polypeptide sequence, including potential motifs, homologous sequences, and methods, algorithms, and searchable databases used for analysis of ORP.

5 Table 3 shows selected fragments of each nucleic acid sequence; the tissue-specific expression patterns of each nucleic acid sequence as determined by northern analysis; diseases, disorders, or conditions associated with these tissues; and the vector into which each cDNA was cloned.

Table 4 describes the tissues used to construct the cDNA libraries from which cDNA clones encoding ORP were isolated.

10 Table 5 shows the tools, programs, and algorithms used to analyze the polynucleotides and polypeptides of the invention, along with applicable descriptions, references, and threshold parameters.

DESCRIPTION OF THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood
15 that this invention is not limited to the particular machines, materials and methods described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a," "an,"
20 and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings
25 as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in
30 connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

DEFINITIONS

"ORP" refers to the amino acid sequences of substantially purified ORP obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and
35 human, and from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term "agonist" refers to a molecule which intensifies or mimics the biological activity of ORP. Agonists may include proteins, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of ORP either by directly interacting with ORP or by acting on components of the biological pathway in which ORP participates.

- 5 An "allelic variant" is an alternative form of the gene encoding ORP. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. A gene may have none, one, or many allelic variants of its naturally occurring form. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides.
- 10 Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

- "Altered" nucleic acid sequences encoding ORP include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as ORP or a polypeptide with at least one functional characteristic of ORP. Included within this definition are
- 15 polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding ORP, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding ORP. The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent ORP. Deliberate
- 20 amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of ORP is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, and positively charged amino acids may include lysine and arginine. Amino acids with uncharged polar side chains having similar hydrophilicity values may
- 25 include: asparagine and glutamine; and serine and threonine. Amino acids with uncharged side chains having similar hydrophilicity values may include: leucine, isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

- The terms "amino acid" and "amino acid sequence" refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic
- 30 molecules. Where "amino acid sequence" is recited to refer to a sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

- "Amplification" relates to the production of additional copies of a nucleic acid sequence. Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known
- 35 in the art.

The term "antagonist" refers to a molecule which inhibits or attenuates the biological activity of ORP. Antagonists may include proteins such as antibodies, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of ORP either by directly interacting with ORP or by acting on components of the biological pathway in which ORP participates.

The term "antibody" refers to intact immunoglobulin molecules as well as to fragments thereof, such as Fab, F(ab')₂, and Fv fragments, which are capable of binding an epitopic determinant. Antibodies that bind ORP polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

The term "antigenic determinant" refers to that region of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (particular regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "antisense" refers to any composition capable of base-pairing with the "sense" (coding) strand of a specific nucleic acid sequence. Antisense compositions may include DNA; RNA; peptide nucleic acid (PNA); oligonucleotides having modified backbone linkages such as phosphorothioates, methylphosphonates, or benzylphosphonates; oligonucleotides having modified sugar groups such as 2'-methoxyethyl sugars or 2'-methoxyethoxy sugars; or oligonucleotides having modified bases such as 5-methyl cytosine, 2'-deoxyuracil, or 7-deaza-2'-deoxyguanosine. Antisense molecules may be produced by any method including chemical synthesis or transcription. Once introduced into a cell, the complementary antisense molecule base-pairs with a naturally occurring nucleic acid sequence produced by the cell to form duplexes which block either transcription or translation. The designation "negative" or "minus" can refer to the antisense strand, and the designation "positive" or "plus" can refer to the sense strand of a reference DNA molecule.

The term "biologically active" refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" or "immunogenic" refers to the capability of the natural, recombinant, or synthetic ORP, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

"Complementary" describes the relationship between two single-stranded nucleic acid

sequences that anneal by base-pairing. For example, 5'-AGT-3' pairs with its complement, 3'-TCA-5'.

A "composition comprising a given polynucleotide sequence" and a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding ORP or fragments of ORP may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"Consensus sequence" refers to a nucleic acid sequence which has been subjected to repeated DNA sequence analysis to resolve uncalled bases, extended using the XL-PCR kit (Applied Biosystems, Foster City CA) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from one or more overlapping cDNA, EST, or genomic DNA fragments using a computer program for fragment assembly, such as the GELVIEW fragment assembly system (GCG, Madison WI) or Phrap (University of Washington, Seattle WA). Some sequences have been both extended and assembled to produce the consensus sequence.

"Conservative amino acid substitutions" are those substitutions that are predicted to least interfere with the properties of the original protein, i.e., the structure and especially the function of the protein is conserved and not significantly changed by such substitutions. The table below shows amino acids which may be substituted for an original amino acid in a protein and which are regarded as conservative amino acid substitutions.

	Original Residue	Conservative Substitution
25	Ala	Gly, Ser
	Arg	His, Lys
	Asn	Asp, Gln, His
	Asp	Asn, Glu
	Cys	Ala, Ser
30	Gln	Asn, Glu, His
	Glu	Asp, Gln, His
	Gly	Ala
	His	Asn, Arg, Gln, Glu
	Ile	Leu, Val
35	Leu	Ile, Val
	Lys	Arg, Gln, Glu
	Met	Leu, Ile
	Phe	His, Met, Leu, Trp, Tyr
	Ser	Cys, Thr
40	Thr	Ser, Val
	Trp	Phe, Tyr

Tyr
Val

His, Phe, Trp
Ile, Leu, Thr

Conservative amino acid substitutions generally maintain (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a beta sheet or alpha helical conformation, (b) the charge or hydrophobicity of the molecule at the site of the substitution, and/or (c) the bulk of the side chain.

A "deletion" refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term "derivative" refers to a chemically modified polynucleotide or polypeptide. Chemical modifications of a polynucleotide sequence can include, for example, replacement of hydrogen by an alkyl, acyl, hydroxyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

A "detectable label" refers to a reporter molecule or enzyme that is capable of generating a measurable signal and is covalently or noncovalently joined to a polynucleotide or polypeptide.

A "fragment" is a unique portion of ORP or the polynucleotide encoding ORP which is identical in sequence to but shorter in length than the parent sequence. A fragment may comprise up to the entire length of the defined sequence, minus one nucleotide/amino acid residue. For example, a fragment may comprise from 5 to 1000 contiguous nucleotides or amino acid residues. A fragment used as a probe, primer, antigen, therapeutic molecule, or for other purposes, may be at least 5, 10, 15, 16, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous nucleotides or amino acid residues in length. Fragments may be preferentially selected from certain regions of a molecule. For example, a polypeptide fragment may comprise a certain length of contiguous amino acids selected from the first 250 or 500 amino acids (or first 25% or 50% of a polypeptide) as shown in a certain defined sequence. Clearly these lengths are exemplary, and any length that is supported by the specification, including the Sequence Listing, tables, and figures, may be encompassed by the present embodiments.

A fragment of SEQ ID NO:28-54 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:28-54, for example, as distinct from any other sequence in the genome from which the fragment was obtained. A fragment of SEQ ID NO:28-54 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:28-54 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:28-54 and the region of SEQ ID NO:28-54 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A fragment of SEQ ID NO:1-27 is encoded by a fragment of SEQ ID NO:28-54. A fragment of SEQ ID NO:1-27 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-27. For example, a fragment of SEQ ID NO:1-27 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-27. The precise length of a fragment of SEQ ID NO:1-27 and the region of SEQ ID NO:1-27 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A "full-length" polynucleotide sequence is one containing at least a translation initiation codon (e.g., methionine) followed by an open reading frame and a translation termination codon. A "full-length" polynucleotide sequence encodes a "full-length" polypeptide sequence.

"Homology" refers to sequence similarity or, interchangeably, sequence identity, between two or more polynucleotide sequences or two or more polypeptide sequences.

The terms "percent identity" and "% identity," as applied to polynucleotide sequences, refer to the percentage of residue matches between at least two polynucleotide sequences aligned using a standardized algorithm. Such an algorithm may insert, in a standardized and reproducible way, gaps in the sequences being compared in order to optimize alignment between two sequences, and therefore achieve a more meaningful comparison of the two sequences.

Percent identity between polynucleotide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989) CABIOS 5:151-153 and in Higgins, D.G. et al. (1992) CABIOS 8:189-191. For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default. Percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polynucleotide sequences.

Alternatively, a suite of commonly used and freely available sequence comparison algorithms is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at <http://www.ncbi.nlm.nih.gov/BLAST/>. The BLAST software suite includes various sequence analysis programs including "blastn," that is used to align a known polynucleotide sequence with other polynucleotide sequences from a variety of databases. Also available is a tool called "BLAST 2 Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 Sequences" can be accessed and used interactively at <http://www.ncbi.nlm.nih.gov/gorf/bl2.html>. The

"BLAST 2 Sequences" tool can be used for both blastn and blastp (discussed below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool Version 2.0.12 (April-21-2000) set at default parameters. Such default parameters may be, for example:

5 *Matrix: BLOSUM62*
 Reward for match: 1
 Penalty for mismatch: -2
 Open Gap: 5 and Extension Gap: 2 penalties
 Gap x drop-off: 50
 10 *Expect: 10*
 Word Size: 11
 Filter: on

Percent identity may be measured over the length of an entire defined sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over
 15 the length of a fragment taken from a larger, defined sequence, for instance, a fragment of at least 20, at least 30, at least 40, at least 50, at least 70, at least 100, or at least 200 contiguous nucleotides. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures, or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

20 Nucleic acid sequences that do not show a high degree of identity may nevertheless encode similar amino acid sequences due to the degeneracy of the genetic code. It is understood that changes in a nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid sequences that all encode substantially the same protein.

The phrases "percent identity" and "% identity," as applied to polypeptide sequences, refer to
 25 the percentage of residue matches between at least two polypeptide sequences aligned using a standardized algorithm. Methods of polypeptide sequence alignment are well-known. Some alignment methods take into account conservative amino acid substitutions. Such conservative substitutions, explained in more detail above, generally preserve the charge and hydrophobicity at the site of substitution, thus preserving the structure (and therefore function) of the polypeptide.

30 Percent identity between polypeptide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program (described and referenced above). For pairwise alignments of polypeptide sequences using CLUSTAL V, the default parameters are set as follows: Ktuple=1, gap penalty=3, window=5, and "diagonals saved"=5. The PAM250 matrix is selected as the default residue weight table. As with
 35 polynucleotide alignments, the percent identity is reported by CLUSTAL V as the "percent similarity".

between aligned polypeptide sequence pairs.

Alternatively the NCBI BLAST software suite may be used. For example, for a pairwise comparison of two polypeptide sequences, one may use the "BLAST 2 Sequences" tool Version 2.0.12 (Apr-21-2000) with blastp set at default parameters. Such default parameters may be, for example:

5 *Matrix: BLOSUM62*
 Open Gap: 11 and Extension Gap: 1 penalties
 Gap x drop-off: 50
 Expect: 10
 Word Size: 3
10 *Filter: on*

Percent identity may be measured over the length of an entire defined polypeptide sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined polypeptide sequence, for instance, a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least 150
15 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

"Human artificial chromosomes" (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size, and which contain all of the elements required for
20 chromosome replication, segregation and maintenance.

The term "humanized antibody" refers to an antibody molecule in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

"Hybridization" refers to the process by which a polynucleotide strand anneals with a
25 complementary strand through base pairing under defined hybridization conditions. Specific hybridization is an indication that two nucleic acid sequences share a high degree of complementarity. Specific hybridization complexes form under permissive annealing conditions and remain hybridized after the "washing" step(s). The washing step(s) is particularly important in determining the stringency of the hybridization process, with more stringent conditions allowing less non-specific binding, i.e.,
30 binding between pairs of nucleic acid strands that are not perfectly matched. Permissive conditions for annealing of nucleic acid sequences are routinely determinable by one of ordinary skill in the art and may be consistent among hybridization experiments, whereas wash conditions may be varied among experiments to achieve the desired stringency, and therefore hybridization specificity. Permissive annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about 1% (w/v)
35 SDS, and about 100 µg/ml sheared, denatured salmon sperm DNA.

Generally, stringency of hybridization is expressed, in part, with reference to the temperature under which the wash step is carried out. Such wash temperatures are typically selected to be about 5°C to 20°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. An equation for calculating T_m and conditions for nucleic acid hybridization are well known and can be found in Sambrook, J. et al., 1989, Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; specifically see volume 2, chapter 9.

High stringency conditions for hybridization between polynucleotides of the present invention include wash conditions of 68°C in the presence of about 0.2 x SSC and about 0.1% SDS, for 1 hour. Alternatively, temperatures of about 65°C, 60°C, 55°C, or 42°C may be used. SSC concentration may be varied from about 0.1 to 2 x SSC, with SDS being present at about 0.1%. Typically, blocking reagents are used to block non-specific hybridization. Such blocking reagents include, for instance, sheared and denatured salmon sperm DNA at about 100-200 µg/ml. Organic solvent, such as formamide at a concentration of about 35-50% v/v, may also be used under particular circumstances, such as for RNA:DNA hybridizations. Useful variations on these wash conditions will be readily apparent to those of ordinary skill in the art. Hybridization, particularly under high stringency conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is strongly indicative of a similar role for the nucleotides and their encoded polypeptides.

The term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C_0t or R_0t analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

The words "insertion" and "addition" refer to changes in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively.

"Immune response" can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect cellular and systemic defense systems.

An "immunogenic fragment" is a polypeptide or oligopeptide fragment of ORP which is capable of eliciting an immune response when introduced into a living organism, for example, a mammal. The term "immunogenic fragment" also includes any polypeptide or oligopeptide fragment of ORP which is useful in any of the antibody production methods disclosed herein or known in the art.

The term "microarray" refers to an arrangement of a plurality of polynucleotides, polypeptides, or other chemical compounds on a substrate.

The terms "element" and "array element" refer to a polynucleotide, polypeptide, or other chemical compound having a unique and defined position on a microarray.

5 The term "modulate" refers to a change in the activity of ORP. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of ORP.

10 The phrases "nucleic acid" and "nucleic acid sequence" refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material.

15 "Operably linked" refers to the situation in which a first nucleic acid sequence is placed in a functional relationship with a second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Operably linked DNA sequences may be in close proximity or contiguous and, where necessary to join two protein coding regions, in the same reading frame.

20 "Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

25 "Post-translational modification" of an ORP may involve lipidation, glycosylation, phosphorylation, acetylation, racemization, proteolytic cleavage, and other modifications known in the art. These processes may occur synthetically or biochemically. Biochemical modifications will vary by cell type depending on the enzymatic milieu of ORP.

30 "Probe" refers to nucleic acid sequences encoding ORP, their complements, or fragments thereof, which are used to detect identical, allelic or related nucleic acid sequences. Probes are isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule. Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes. "Primers" are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target polynucleotide by complementary base-pairing. The primer may then be extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification (and identification) of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR).

35 Probes and primers as used in the present invention typically comprise at least 15 contiguous nucleotides of a known sequence. In order to enhance specificity, longer probes and primers may also

be employed, such as probes and primers that comprise at least 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, or at least 150 consecutive nucleotides of the disclosed nucleic acid sequences. Probes and primers may be considerably longer than these examples, and it is understood that any length supported by the specification, including the tables, figures, and Sequence Listing, may be used.

5 Methods for preparing and using probes and primers are described in the references, for example Sambrook, J. et al. (1989) Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; Ausubel, F.M. et al. (1987) Current Protocols in Molecular Biology, Greene Publ. Assoc. & Wiley-Intersciences, New York NY; Innis, M. et al. (1990) PCR Protocols, A Guide to Methods and Applications, Academic Press, San Diego CA. PCR primer pairs
10 can be derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA).

 Oligonucleotides for use as primers are selected using software known in the art for such purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to
15 100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to 5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer selection programs have incorporated additional features for expanded capabilities. For example, the PrimOU primer selection program (available to the public from the Genome Center at University of Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from megabase sequences
20 and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for
25 microarrays. (The source code for the latter two primer selection programs may also be obtained from their respective sources and modified to meet the user's specific needs.) The PrimeGen program
30 (available to the public from the UK Human Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments, thereby allowing selection of primers that hybridize to either the most conserved or least conserved regions of aligned nucleic acid sequences.
 Hence, this program is useful for identification of both unique and conserved oligonucleotides and
30 polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the
 above selection methods are useful in hybridization technologies, for example, as PCR or sequencing primers, microarray elements, or specific probes to identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

35 A "recombinant nucleic acid" is a sequence that is not naturally occurring or has a sequence

that is made by an artificial combination of two or more otherwise separated segments of sequence. This artificial combination is often accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques such as those described in Sambrook, *supra*. The term recombinant includes nucleic acids that have
5 been altered solely by addition, substitution, or deletion of a portion of the nucleic acid. Frequently, a recombinant nucleic acid may include a nucleic acid sequence operably linked to a promoter sequence. Such a recombinant nucleic acid may be part of a vector that is used, for example, to transform a cell.

Alternatively, such recombinant nucleic acids may be part of a viral vector, e.g., based on a vaccinia virus, that could be used to vaccinate a mammal wherein the recombinant nucleic acid is
10 expressed, inducing a protective immunological response in the mammal.

A "regulatory element" refers to a nucleic acid sequence usually derived from untranslated regions of a gene and includes enhancers, promoters, introns, and 5' and 3' untranslated regions (UTRs). Regulatory elements interact with host or viral proteins which control transcription, translation, or RNA stability.

15 "Reporter molecules" are chemical or biochemical moieties used for labeling a nucleic acid, amino acid, or antibody. Reporter molecules include radionuclides; enzymes; fluorescent, chemiluminescent, or chromogenic agents; substrates; cofactors; inhibitors; magnetic particles; and other moieties known in the art.

An "RNA equivalent," in reference to a DNA sequence, is composed of the same linear
20 sequence of nucleotides as the reference DNA sequence with the exception that all occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The term "sample" is used in its broadest sense. A sample suspected of containing nucleic acids encoding ORP, or fragments thereof, or ORP itself, may comprise a bodily fluid; an extract from
25 a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

The terms "specific binding" and "specifically binding" refer to that interaction between a protein or peptide and an agonist, an antibody, an antagonist, a small molecule, or any natural or synthetic binding composition. The interaction is dependent upon the presence of a particular structure
30 of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide comprising the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

The term "substantially purified" refers to nucleic acid or amino acid sequences that are
35 removed from their natural environment and are isolated or separated, and are at least 60% free,

preferably at least 75% free, and most preferably at least 90% free from other components with which they are naturally associated.

A "substitution" refers to the replacement of one or more amino acid residues or nucleotides by different amino acid residues or nucleotides, respectively.

5 "Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

A "transcript image" refers to the collective pattern of gene expression by a particular cell type
10 or tissue under given conditions at a given time.

"Transformation" describes a process by which exogenous DNA is introduced into a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type
15 of host cell being transformed and may include, but is not limited to, bacteriophage or viral infection, electroporation, heat shock, lipofection, and particle bombardment. The term "transformed" cells includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

20 A "transgenic organism," as used herein, is any organism, including but not limited to animals and plants, in which one or more of the cells of the organism contains heterologous nucleic acid introduced by way of human intervention, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with
25 a recombinant virus. The term genetic manipulation does not include classical cross-breeding, or in vitro fertilization, but rather is directed to the introduction of a recombinant DNA molecule. The transgenic organisms contemplated in accordance with the present invention include bacteria, cyanobacteria, fungi, plants, and animals. The isolated DNA of the present invention can be introduced into the host by methods known in the art, for example infection, transfection,
30 transformation or transconjugation. Techniques for transferring the DNA of the present invention into such organisms are widely known and provided in references such as Sambrook, J. et al. (1989), supra.

A "variant" of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 40% sequence identity to the particular nucleic acid sequence over a certain length of one of the
35 nucleic acid sequences using blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999)

set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 95% or at least 98% or greater sequence identity over a certain defined length. A variant may be described as, for example, an "allelic" (as defined above), "splice," "species," or "polymorphic" variant. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternative splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or lack domains that are present in the reference molecule. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides generally will have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

A "variant" of a particular polypeptide sequence is defined as a polypeptide sequence having at least 40% sequence identity to the particular polypeptide sequence over a certain length of one of the polypeptide sequences using blastp with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of polypeptides may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or at least 98% or greater sequence identity over a certain defined length of one of the polypeptides.

THE INVENTION

The invention is based on the discovery of new human oxidoreductase proteins (ORP), the polynucleotides encoding ORP, and the use of these compositions for the diagnosis, treatment, or prevention of cell proliferative disorders including cancer; endocrine, metabolic, reproductive, neurological, viral, and autoimmune/inflammatory disorders.

Table 1 lists the Incyte clones used to assemble full length nucleotide sequences encoding ORP. Columns 1 and 2 show the sequence identification numbers (SEQ ID NOs) of the polypeptide and nucleotide sequences, respectively. Column 3 shows the clone IDs of the Incyte clones in which nucleic acids encoding each ORP were identified, and column 4 shows the cDNA libraries from which these clones were isolated. Column 5 shows Incyte clones and their corresponding cDNA libraries. Clones for which cDNA libraries are not indicated were derived from pooled cDNA libraries. In some cases, GenBank sequence identifiers are also shown in column 5. The Incyte clones and GenBank cDNA sequences, where indicated, in column 5 were used to assemble the consensus nucleotide sequence of each ORP and are useful as fragments in hybridization technologies.

The columns of Table 2 show various properties of each of the polypeptides of the invention: column 1 references the SEQ ID NO; column 2 shows the number of amino acid residues in each

polypeptide; column 3 shows potential phosphorylation sites; column 4 shows potential glycosylation sites; column 5 shows the amino acid residues comprising signature sequences and motifs; column 6 shows homologous sequences as identified by BLAST analysis; and column 7 shows analytical methods and in some cases, searchable databases to which the analytical methods were applied. The methods of column 7 were used to characterize each polypeptide through sequence homology and protein motifs.

The columns of Table 3 show the tissue-specificity and diseases, disorders, or conditions associated with nucleotide sequences encoding ORP. The first column of Table 3 lists the nucleotide SEQ ID NOs. Column 2 lists fragments of the nucleotide sequences of column 1. These fragments are useful, for example, in hybridization or amplification technologies to identify SEQ ID NO:28-54 and to distinguish between SEQ ID NO:28-54 and related polynucleotide sequences. The polypeptides encoded by these fragments are useful, for example, as immunogenic peptides. Column 3 lists tissue categories which express ORP as a fraction of total tissues expressing ORP. Column 4 lists diseases, disorders, or conditions associated with those tissues expressing ORP as a fraction of total tissues expressing ORP. Column 5 lists the vectors used to subclone each cDNA library.

The columns of Table 4 show descriptions of the tissues used to construct the cDNA libraries from which cDNA clones encoding ORP were isolated. Column 1 references the nucleotide SEQ ID NOs, column 2 shows the cDNA libraries from which these clones were isolated, and column 3 shows the tissue origins and other descriptive information relevant to the cDNA libraries in column 2.

SEQ ID NO:31 maps to chromosome 3 within the interval from 23.2 to 31.4 centiMorgans. SEQ ID NO:42 maps to chromosome 22 within the interval from 0 to 40.2 centiMorgans. SEQ ID NO:48 maps to chromosome 7 within the interval from 100.5 to 114.5 centiMorgans, to chromosome 17 within the interval from 67.6 to 69.3 centiMorgans, and to chromosome 17 within the interval from 83.8 centiMorgans to the q terminus. SEQ ID NO:53 maps to chromosome 11 within the interval from 64.9 to 70.9 centiMorgans.

The invention also encompasses ORP variants. A preferred ORP variant is one which has at least about 80%, or alternatively at least about 90%, or even at least about 95% amino acid sequence identity to the ORP amino acid sequence, and which contains at least one functional or structural characteristic of ORP.

The invention also encompasses polynucleotides which encode ORP. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:28-54, which encodes ORP. The polynucleotide sequences of SEQ ID NO:28-54, as presented in the Sequence Listing, embrace the equivalent RNA sequences, wherein occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The invention also encompasses a variant of a polynucleotide sequence encoding ORP. In

particular, such a variant polynucleotide sequence will have at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding ORP. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:28-54 which has at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:28-54. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of ORP.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding ORP, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring ORP, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode ORP and its variants are generally capable of hybridizing to the nucleotide sequence of the naturally occurring ORP under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding ORP or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding ORP and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode ORP and ORP derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding ORP or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:28-54 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) *Methods Enzymol.* 152:399-407; Kimmel, A.R. (1987) *Methods Enzymol.* 152:507-511.) Hybridization conditions, including annealing and wash conditions, are described in

"Definitions."

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Applied Biosystems, Foster City CA), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), PTC200 thermal cycler (MJ Research, Watertown MA) and ABI CATALYST 800 thermal cycler (Applied Biosystems). Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing system (Applied Biosystems), the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

The nucleic acid sequences encoding ORP may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) PCR Methods Applic. 2:318-322.) Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) Nucleic Acids Res. 19:3055-3060).

Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 Primer Analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to

72°C.

When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library
5 does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-
10 specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Applied Biosystems), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments
15 which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode ORP may be cloned in recombinant DNA molecules that direct expression of ORP, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent
20 amino acid sequence may be produced and used to express ORP.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter ORP-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic
25 oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

The nucleotides of the present invention may be subjected to DNA shuffling techniques such as MOLECULARBREEDING (Maxygen Inc., Santa Clara CA; described in U.S. Patent Number
30 5,837,458; Chang, C.-C. et al. (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. et al. (1999) Nat. Biotechnol. 17:259-264; and Crameri, A. et al. (1996) Nat. Biotechnol. 14:315-319) to alter or improve the biological properties of ORP, such as its biological or enzymatic activity or its ability to bind to other molecules or compounds. DNA shuffling is a process by which a library of gene variants is produced using PCR-mediated recombination of gene fragments. The library is then
35 subjected to selection or screening procedures that identify those gene variants with the desired

properties. These preferred variants may then be pooled and further subjected to recursive rounds of DNA shuffling and selection/screening. Thus, genetic diversity is created through "artificial" breeding and rapid molecular evolution. For example, fragments of a single gene containing random point mutations may be recombined, screened, and then reshuffled until the desired properties are optimized. Alternatively, fragments of a given gene may be recombined with fragments of homologous genes in the same gene family, either from the same or different species, thereby maximizing the genetic diversity of multiple naturally occurring genes in a directed and controllable manner.

In another embodiment, sequences encoding ORP may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) *Nucleic Acids Symp. Ser.* 7:215-223; Horn, T. et al. (1980) *Nucleic Acids Symp. Ser.* 7:225-232.) Alternatively, ORP itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solution-phase or solid-phase techniques. (See, e.g., Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY, pp. 55-60; and Roberge, J.Y. et al. (1995) *Science* 269:202-204.) Automated synthesis may be achieved using the ABI 431A peptide synthesizer (Applied Biosystems). Additionally, the amino acid sequence of ORP, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide or a polypeptide having a sequence of a naturally occurring polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography. (See, e.g., Chiez, R.M. and F.Z. Regnier (1990) *Methods Enzymol.* 182:392-421.) The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing. (See, e.g., Creighton, supra, pp. 28-53.)

In order to express a biologically active ORP, the nucleotide sequences encoding ORP or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding ORP. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding ORP. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding ORP and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the

vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994) *Results Probl. Cell Differ.* 20:125-162.)

- 5 Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding ORP and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. (See, e.g., Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

- 10 A variety of expression vector/host systems may be utilized to contain and express sequences encoding ORP. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus);
- 15 plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. (See, e.g., Sambrook, supra; Ausubel, supra; Van Heeke, G. and S.M. Schuster (1989) *J. Biol. Chem.* 264:5503-5509; Bitter, G.A. et al. (1987) *Methods Enzymol.* 153:516-544; Scorer, C.A. et al. (1994) *Bio/Technology* 12:181-184; Engelhard, E.K. et al. (1994) *Proc. Natl.*
- 20 *Acad. Sci. USA* 91:3224-3227; Sandig, V. et al. (1996) *Hum. Gene Ther.* 7:1937-1945; Takamatsu, N. (1987) *EMBO J.* 6:307-311; Coruzzi, G. et al. (1984) *EMBO J.* 3:1671-1680; Broglie, R. et al. (1984) *Science* 224:838-843; Winter, J. et al. (1991) *Results Probl. Cell Differ.* 17:85-105; The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196; Logan, J. and T. Shenk (1984) *Proc. Natl. Acad. Sci. USA* 81:3655-3659; and Harrington, J.J. et al. (1997) *Nat. Genet.* 15:345-355.) Expression vectors derived from retroviruses,
- 25 adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. (See, e.g., Di Nicola, M. et al. (1998) *Cancer Gen. Ther.* 5(6):350-356; Yu, M. et al. (1993) *Proc. Natl. Acad. Sci. USA* 90(13):6340-6344; Buller, R.M. et al. (1985) *Nature* 317(6040):813-815; McGregor, D.P. et al.
- 30 (1994) *Mol. Immunol.* 31(3):219-226; and Verma, I.M. and N. Somia (1997) *Nature* 389:239-242.) The invention is not limited by the host cell employed.

- In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding ORP. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding ORP can be achieved using a
- 35 multifunctional E. coli vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPORT1 plasmid

(Life Technologies). Ligation of sequences encoding ORP into the vector's multiple cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for in vitro transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) *J. Biol. Chem.* 264:5503-5509.) When large quantities of ORP are needed, e.g. for the production of antibodies, vectors which direct high level expression of ORP may be used. For example, vectors containing the strong, inducible T5 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of ORP. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast *Saccharomyces cerevisiae* or *Pichia pastoris*. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, supra; Bitter, supra; and Scorer, supra.)

Plant systems may also be used for expression of ORP. Transcription of sequences encoding ORP may be driven viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) *EMBO J.* 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, supra; Broglie, supra; and Winter, supra.) These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding ORP may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses ORP in host cells. (See, e.g., Logan, J. and T. Shenk (1984) *Proc. Natl. Acad. Sci. USA* 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) *Nat. Genet.* 15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression of

ORP in cell lines is preferred. For example, sequences encoding ORP can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being
5 switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include,
10 but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk* and *ap^r* cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to
15 chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), β
20 glucuronidase and its substrate β -glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is
25 also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding ORP is inserted within a marker gene sequence, transformed cells containing sequences encoding ORP can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding ORP under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates
30 expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding ORP and that express ORP may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based
35 technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of ORP using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on ORP is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ.)

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding ORP include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding ORP, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Amersham Pharmacia Biotech, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding ORP may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode ORP may be designed to contain signal sequences which direct secretion of ORP through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" or "pro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing

of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding ORP may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric ORP protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of ORP activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the ORP encoding sequence and the heterologous protein sequence, so that ORP may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, supra, ch. 10). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

In a further embodiment of the invention, synthesis of radiolabeled ORP may be achieved in vitro using the TNT rabbit reticulocyte lysate or wheat germ extract system (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, for example, ³⁵S-methionine.

ORP of the present invention or fragments thereof may be used to screen for compounds that specifically bind to ORP. At least one and up to a plurality of test compounds may be screened for specific binding to ORP. Examples of test compounds include antibodies, oligonucleotides, proteins (e.g., receptors), or small molecules.

In one embodiment, the compound thus identified is closely related to the natural ligand of ORP, e.g., a ligand or fragment thereof, a natural substrate, a structural or functional mimetic, or a natural binding partner. (See, e.g., Coligan, J.E. et al. (1991) Current Protocols in Immunology 1(2): Chapter 5.) Similarly, the compound can be closely related to the natural receptor to which ORP binds, or to at least a fragment of the receptor, e.g., the ligand binding site. In either case, the compound can be rationally designed using known techniques. In one embodiment, screening for these compounds involves producing appropriate cells which express ORP, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, Drosophila, or E.

coli. Cells expressing ORP or cell membrane fractions which contain ORP are then contacted with a test compound and binding, stimulation, or inhibition of activity of either ORP or the compound is analyzed.

An assay may simply test binding of a test compound to the polypeptide, wherein binding is
5 detected by a fluorophore, radioisotope, enzyme conjugate, or other detectable label. For example, the assay may comprise the steps of combining at least one test compound with ORP, either in solution or affixed to a solid support, and detecting the binding of ORP to the compound. Alternatively, the assay may detect or measure binding of a test compound in the presence of a labeled competitor. Additionally, the assay may be carried out using cell-free preparations, chemical
10 libraries, or natural product mixtures, and the test compound(s) may be free in solution or affixed to a solid support.

ORP of the present invention or fragments thereof may be used to screen for compounds that modulate the activity of ORP. Such compounds may include agonists, antagonists, or partial or inverse agonists. In one embodiment, an assay is performed under conditions permissive for ORP
15 activity, wherein ORP is combined with at least one test compound, and the activity of ORP in the presence of a test compound is compared with the activity of ORP in the absence of the test compound. A change in the activity of ORP in the presence of the test compound is indicative of a compound that modulates the activity of ORP. Alternatively, a test compound is combined with an in vitro or cell-free system comprising ORP under conditions suitable for ORP activity, and the assay is
20 performed. In either of these assays, a test compound which modulates the activity of ORP may do so indirectly and need not come in direct contact with the test compound. At least one and up to a plurality of test compounds may be screened.

In another embodiment, polynucleotides encoding ORP or their mammalian homologs may be "knocked out" in an animal model system using homologous recombination in embryonic stem
25 (ES) cells. Such techniques are well known in the art and are useful for the generation of animal models of human disease. (See, e.g., U.S. Patent No. 5,175,383 and U.S. Patent No. 5,767,337.) For example, mouse ES cells, such as the mouse 129/SvJ cell line, are derived from the early mouse embryo and grown in culture. The ES cells are transformed with a vector containing the gene of interest disrupted by a marker gene, e.g., the neomycin phosphotransferase gene (neo; Capecchi, M.R.
30 (1989) Science 244:1288-1292). The vector integrates into the corresponding region of the host genome by homologous recombination. Alternatively, homologous recombination takes place using the Cre-loxP system to knockout a gene of interest in a tissue- or developmental stage-specific manner (Marth, J.D. (1996) Clin. Invest. 97:1999-2002; Wagner, K.U. et al. (1997) Nucleic Acids Res. 25:4323-4330). Transformed ES cells are identified and microinjected into mouse cell
35 blastocysts such as those from the C57BL/6 mouse strain. The blastocysts are surgically transferred

to pseudopregnant dams, and the resulting chimeric progeny are genotyped and bred to produce heterozygous or homozygous strains. Transgenic animals thus generated may be tested with potential therapeutic or toxic agents.

Polynucleotides encoding ORP may also be manipulated in vitro in ES cells derived from human blastocysts. Human ES cells have the potential to differentiate into at least eight separate cell lineages including endoderm, mesoderm, and ectodermal cell types. These cell lineages differentiate into, for example, neural cells, hematopoietic lineages, and cardiomyocytes (Thomson, J.A. et al. (1998) Science 282:1145-1147).

Polynucleotides encoding ORP can also be used to create "knockin" humanized animals (pigs) or transgenic animals (mice or rats) to model human disease. With knockin technology, a region of a polynucleotide encoding ORP is injected into animal ES cells, and the injected sequence integrates into the animal cell genome. Transformed cells are injected into blastulae, and the blastulae are implanted as described above. Transgenic progeny or inbred lines are studied and treated with potential pharmaceutical agents to obtain information on treatment of a human disease. Alternatively, a mammal inbred to overexpress ORP, e.g., by secreting ORP in its milk, may also serve as a convenient source of that protein (Janne, J. et al. (1998) Biotechnol. Annu. Rev. 4:55-74).

THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of ORP and oxidoreductase proteins. In addition, the expression of ORP is closely associated with cancerous, inflamed, reproductive, and gastrointestinal tissues. Therefore, ORP appears to play a role in cell proliferative disorders including cancer; endocrine, metabolic, reproductive, neurological, viral, and autoimmune/inflammatory disorders. In the treatment of disorders associated with increased ORP expression or activity, it is desirable to decrease the expression or activity of ORP. In the treatment of disorders associated with decreased ORP expression or activity, it is desirable to increase the expression or activity of ORP.

Therefore, in one embodiment, ORP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of ORP. Examples of such disorders include, but are not limited to, a cell proliferative disorder, such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease, myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia; a cancer, such as adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an endocrine disorder, such as disorders of the hypothalamus and pituitary

resulting from lesions such as primary brain tumors, adenomas, infarction associated with pregnancy, hypophysectomy, aneurysms, vascular malformations, thrombosis, infections, immunological disorders, and complications due to head trauma, disorders associated with hypopituitarism including hypogonadism, Sheehan syndrome, diabetes insipidus, Kallman's disease, Hand-Schuller-Christian disease, Letterer-Siwe disease, sarcoidosis, empty sella syndrome, and dwarfism, disorders associated with hyperpituitarism including acromegaly, gigantism, and syndrome of inappropriate antidiuretic hormone (ADH) secretion (SIADH) often caused by benign adenoma, disorders associated with hypothyroidism including goiter, myxedema, acute thyroiditis associated with bacterial infection, subacute thyroiditis associated with viral infection, autoimmune thyroiditis (Hashimoto's disease), and cretinism, disorders associated with hyperthyroidism including thyrotoxicosis and its various forms, Grave's disease, pretibial myxedema, toxic multinodular goiter, thyroid carcinoma, and Plummer's disease, disorders associated with hyperparathyroidism including Conn disease (chronic hypercalcemia), pancreatic disorders such as Type I or Type II diabetes mellitus and associated complications, disorders associated with the adrenals such as hyperplasia, carcinoma, or adenoma of the adrenal cortex, hypertension associated with alkalosis, amyloidosis, hypokalemia, Cushing's disease, Liddle's syndrome, and Arnold-Healy-Gordon syndrome, pheochromocytoma tumors, and Addison's disease, disorders associated with gonadal steroid hormones such as: in women, abnormal prolactin production, infertility, endometriosis, perturbations of the menstrual cycle, polycystic ovarian disease, hyperprolactinemia, isolated gonadotropin deficiency, amenorrhea, galactorrhea, hermaphroditism, hirsutism and virilization, breast cancer, and, in post-menopausal women, osteoporosis, and, in men, Leydig cell deficiency, male climacteric phase, and germinal cell aplasia, hypergonadal disorders associated with Leydig cell tumors, androgen resistance associated with absence of androgen receptors, syndrome of 5 α -reductase, and gynecomastia; a metabolic disorder, such as Addison's disease, cystic fibrosis, diabetes, fatty hepatocirrhosis, galactosemia, goiter, hyperadrenalism, hypoadrenalism, hyperparathyroidism, hypoparathyroidism, hypercholesterolemia, hyperthyroidism, hypothyroidism hyperlipidemia, hyperlipemia, lipid myopathies, obesity, lipodystrophies, and phenylketonuria, congenital adrenal hyperplasia, pseudovitamin D-deficiency rickets, cerebrotendinous xanthomatosis, and coumarin resistance; a reproductive disorder, such as disorders of prolactin production, infertility, including tubal disease, ovulatory defects, and endometriosis, disruptions of the estrous cycle, disruptions of the menstrual cycle, polycystic ovary syndrome, ovarian hyperstimulation syndrome, endometrial and ovarian tumors, uterine fibroids, autoimmune disorders, ectopic pregnancies, and teratogenesis; cancer of the breast, fibrocystic breast disease, and galactorrhea, disruptions of spermatogenesis, abnormal sperm physiology, cancer of the testis, cancer of the prostate, benign prostatic hyperplasia, prostatitis, Peyronie's disease, impotence, carcinoma of the male breast, and gynecomastia; a neurological disorder, such as epilepsy, ischemic

cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathisia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, and Tourette's disorder; an autoimmune/inflammatory disorder, such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; and a viral disorder, such as viral infections, e.g., those caused by adenoviruses (acute respiratory disease, pneumonia), arenaviruses (lymphocytic choriomeningitis), bunyaviruses (Hantavirus), coronaviruses (pneumonia, chronic bronchitis), hepadnaviruses (hepatitis), herpesviruses (herpes simplex virus, varicella-zoster virus, Epstein-Barr virus, cytomegalovirus), flaviviruses (yellow fever), orthomyxoviruses (influenza), papillomaviruses (cancer), paramyxoviruses (measles, mumps), picornaviruses (rhinovirus, poliovirus, coxsackie-virus),

polyomaviruses (BK virus, JC virus), poxviruses (smallpox), reovirus (Colorado tick fever), retroviruses (human immunodeficiency virus, human T lymphotropic virus), rhabdoviruses (rabies), rotaviruses (gastroenteritis), and togaviruses (encephalitis, rubella).

In another embodiment, a vector capable of expressing ORP or a fragment or derivative thereof
5 may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of ORP including, but not limited to, those described above.

In a further embodiment, a composition comprising a substantially purified ORP in conjunction
10 with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of ORP including, but not limited to, those provided above.

In still another embodiment, an agonist which modulates the activity of ORP may be
administered to a subject to treat or prevent a disorder associated with decreased expression or activity
of ORP including, but not limited to, those listed above.

In a further embodiment, an antagonist of ORP may be administered to a subject to treat or
15 prevent a disorder associated with increased expression or activity of ORP. Examples of such disorders include, but are not limited to, those cell proliferative disorders including cancer; endocrine, metabolic, reproductive, neurological, viral, and autoimmune/inflammatory disorders described above. In one aspect, an antibody which specifically binds ORP may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissues
20 which express ORP.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding ORP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of ORP including, but not limited to, those described above.

In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary
25 sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with
30 lower dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of ORP may be produced using methods which are generally known in the art. In particular, purified ORP may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind ORP. Antibodies to ORP may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal,
35 monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab

expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are generally preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with ORP or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to ORP have an amino acid sequence consisting of at least about 5 amino acids, and generally will consist of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein. Short stretches of ORP amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to ORP may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler, G. et al. (1975) *Nature* 256:495-497; Kozbor, D. et al. (1985) *J. Immunol. Methods* 81:31-42; Cote, R.J. et al. (1983) *Proc. Natl. Acad. Sci. USA* 80:2026-2030; and Cole, S.P. et al. (1984) *Mol. Cell Biol.* 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) *Proc. Natl. Acad. Sci. USA* 81:6851-6855; Neuberger, M.S. et al. (1984) *Nature* 312:604-608; and Takeda, S. et al. (1985) *Nature* 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce ORP-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton, D.R. (1991) *Proc. Natl. Acad. Sci. USA* 88:10134-10137.)

Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) *Proc. Natl. Acad. Sci. USA* 86:3833-3837; Winter, G. et al. (1991) *Nature* 349:293-299.)

Antibody fragments which contain specific binding sites for ORP may also be generated. For

example, such fragments include, but are not limited to, $F(ab')_2$ fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the $F(ab')_2$ fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) Science 246:1275-1281.)

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between ORP and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering ORP epitopes is generally used, but a competitive binding assay may also be employed (Pound, *supra*).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for ORP. Affinity is expressed as an association constant, K_a , which is defined as the molar concentration of ORP-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The K_a determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple ORP epitopes, represents the average affinity, or avidity, of the antibodies for ORP. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular ORP epitope, represents a true measure of affinity. High-affinity antibody preparations with K_a ranging from about 10^9 to 10^{12} L/mole are preferred for use in immunoassays in which the ORP-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with K_a ranging from about 10^6 to 10^7 L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of ORP, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington DC; Liddell, J.E. and A. Cryer (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation of ORP-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available. (See, e.g., Catty, *supra*, and Coligan et al., *supra*.)

In another embodiment of the invention, the polynucleotides encoding ORP, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, modifications of gene

expression can be achieved by designing complementary sequences or antisense molecules (DNA, RNA, PNA, or modified oligonucleotides) to the coding or regulatory regions of the gene encoding ORP.

Such technology is well known in the art, and antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding ORP. (See,

5 e.g., Agrawal, S., ed. (1996) Antisense Therapeutics, Humana Press Inc., Totawa NJ.)

In therapeutic use, any gene delivery system suitable for introduction of the antisense sequences into appropriate target cells can be used. Antisense sequences can be delivered

intracellularly in the form of an expression plasmid which, upon transcription, produces a sequence complementary to at least a portion of the cellular sequence encoding the target protein. (See, e.g.,

10 Slater, J.E. et al. (1998) *J. Allergy Clin. Immunol.* 102(3):469-475; and Scanlon, K.J. et al. (1995) 9(13):1288-1296.) Antisense sequences can also be introduced intracellularly through the use of viral vectors, such as retrovirus and adeno-associated virus vectors. (See, e.g., Miller, A.D. (1990) *Blood* 76:271; Ausubel, *supra*; Uckert, W. and W. Walther (1994) *Pharmacol. Ther.* 63(3):323-347.) Other gene delivery mechanisms include liposome-derived systems, artificial viral envelopes, and other

15 systems known in the art. (See, e.g., Rossi, J.J. (1995) *Br. Med. Bull.* 51(1):217-225; Boado, R.J. et al. (1998) *J. Pharm. Sci.* 87(11):1308-1315; and Morris, M.C. et al. (1997) *Nucleic Acids Res.* 25(14):2730-2736.)

In another embodiment of the invention, polynucleotides encoding ORP may be used for somatic or germline gene therapy. Gene therapy may be performed to (i) correct a genetic deficiency

20 (e.g., in the cases of severe combined immunodeficiency (SCID)-X1 disease characterized by X-linked inheritance (Cavazzana-Calvo, M. et al. (2000) *Science* 288:669-672), severe combined immunodeficiency syndrome associated with an inherited adenosine deaminase (ADA) deficiency

(Blaese, R.M. et al. (1995) *Science* 270:475-480; Bordignon, C. et al. (1995) *Science* 270:470-475), cystic fibrosis (Zabner, J. et al. (1993) *Cell* 75:207-216; Crystal, R.G. et al. (1995) *Hum. Gene*

25 *Therapy* 6:643-666; Crystal, R.G. et al. (1995) *Hum. Gene Therapy* 6:667-703), thalassemias, familial hypercholesterolemia, and hemophilia resulting from Factor VIII or Factor IX deficiencies (Crystal, R.G. (1995) *Science* 270:404-410; Verma, I.M. and N. Somia (1997) *Nature* 389:239-242)), (ii)

express a conditionally lethal gene product (e.g., in the case of cancers which result from unregulated cell proliferation), or (iii) express a protein which affords protection against intracellular parasites (e.g.,

30 against human retroviruses, such as human immunodeficiency virus (HIV) (Baltimore, D. (1988) *Nature* 335:395-396; Poeschla, E. et al. (1996) *Proc. Natl. Acad. Sci. USA.* 93:11395-11399),

hepatitis B or C virus (HBV, HCV); fungal parasites, such as Candida albicans and Paracoccidioides brasiliensis; and protozoan parasites such as Plasmodium falciparum and Trypanosoma cruzi). In the

case where a genetic deficiency in ORP expression or regulation causes disease, the expression of ORP

35 from an appropriate population of transduced cells may alleviate the clinical manifestations caused by

the genetic deficiency.

In a further embodiment of the invention, diseases or disorders caused by deficiencies in ORP are treated by constructing mammalian expression vectors encoding ORP and introducing these vectors by mechanical means into ORP-deficient cells. Mechanical transfer technologies for use with cells in vivo or ex vitro include (i) direct DNA microinjection into individual cells, (ii) ballistic gold particle delivery, (iii) liposome-mediated transfection, (iv) receptor-mediated gene transfer, and (v) the use of DNA transposons (Morgan, R.A. and W.F. Anderson (1993) *Annu. Rev. Biochem.* 62:191-217; Ivics, Z. (1997) *Cell* 91:501-510; Boulay, J-L. and H. Récipon (1998) *Curr. Opin. Biotechnol.* 9:445-450).

Expression vectors that may be effective for the expression of ORP include, but are not limited to, the pCDNA 3.1, EPITAG, PRCCMV2, PREP, PVAX vectors (Invitrogen, Carlsbad CA), PCMV-SCRIPT, PCMV-TAG, PEGSH/PERV (Stratagene, La Jolla CA), and PTET-OFF, PTET-ON, PTRE2, PTRE2-LUC, PTK-HYG (Clontech, Palo Alto CA). ORP may be expressed using (i) a constitutively active promoter, (e.g., from cytomegalovirus (CMV), Rous sarcoma virus (RSV), SV40 virus, thymidine kinase (TK), or β -actin genes), (ii) an inducible promoter (e.g., the tetracycline-regulated promoter (Gossen, M. and H. Bujard (1992) *Proc. Natl. Acad. Sci. USA* 89:5547-5551; Gossen, M. et al. (1995) *Science* 268:1766-1769; Rossi, F.M.V. and H.M. Blau (1998) *Curr. Opin. Biotechnol.* 9:451-456), commercially available in the T-REX plasmid (Invitrogen)); the ecdysone-inducible promoter (available in the plasmids PVGRXR and PIND; Invitrogen); the FK506/rapamycin inducible promoter; or the RU486/mifepristone inducible promoter (Rossi, F.M.V. and H.M. Blau, *supra*)), or (iii) a tissue-specific promoter or the native promoter of the endogenous gene encoding ORP from a normal individual.

Commercially available liposome transformation kits (e.g., the PERFECT LIPID TRANSFECTION KIT, available from Invitrogen) allow one with ordinary skill in the art to deliver polynucleotides to target cells in culture and require minimal effort to optimize experimental parameters. In the alternative, transformation is performed using the calcium phosphate method (Graham, F.L. and A.J. Eb (1973) *Virology* 52:456-467), or by electroporation (Neumann, E. et al. (1982) *EMBO J.* 1:841-845). The introduction of DNA to primary cells requires modification of these standardized mammalian transfection protocols.

In another embodiment of the invention, diseases or disorders caused by genetic defects with respect to ORP expression are treated by constructing a retrovirus vector consisting of (i) the polynucleotide encoding ORP under the control of an independent promoter or the retrovirus long terminal repeat (LTR) promoter, (ii) appropriate RNA packaging signals, and (iii) a Rev-responsive element (RRE) along with additional retrovirus *cis*-acting RNA sequences and coding sequences required for efficient vector propagation. Retrovirus vectors (e.g., PFB and PFBNEO) are commercially available (Stratagene) and are based on published data (Riviere, I. et al. (1995) *Proc.*

Natl. Acad. Sci. USA 92:6733-6737), incorporated by reference herein. The vector is propagated in an appropriate vector producing cell line (VPCL) that expresses an envelope gene with a tropism for receptors on the target cells or a promiscuous envelope protein such as VSVg (Armentano, D. et al. (1987) J. Virol. 61:1647-1650; Bender, M.A. et al. (1987) J. Virol. 61:1639-1646; Adam, M.A. and A.D. Miller (1988) J. Virol. 62:3802-3806; Dull, T. et al. (1998) J. Virol. 72:8463-8471; Zufferey, R. et al. (1998) J. Virol. 72:9873-9880). U.S. Patent Number 5,910,434 to Rigg ("Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant") discloses a method for obtaining retrovirus packaging cell lines and is hereby incorporated by reference.

Propagation of retrovirus vectors, transduction of a population of cells (e.g., CD4⁺ T-cells), and the return of transduced cells to a patient are procedures well known to persons skilled in the art of gene therapy and have been well documented (Ranga, U. et al. (1997) J. Virol. 71:7020-7029; Bauer, G. et al. (1997) Blood 89:2259-2267; Bonyhadi, M.L. (1997) J. Virol. 71:4707-4716; Ranga, U. et al. (1998) Proc. Natl. Acad. Sci. USA 95:1201-1206; Su, L. (1997) Blood 89:2283-2290).

In the alternative, an adenovirus-based gene therapy delivery system is used to deliver polynucleotides encoding ORP to cells which have one or more genetic abnormalities with respect to the expression of ORP. The construction and packaging of adenovirus-based vectors are well known to those with ordinary skill in the art. Replication defective adenovirus vectors have proven to be versatile for importing genes encoding immunoregulatory proteins into intact islets in the pancreas (Csete, M.E. et al. (1995) Transplantation 27:263-268). Potentially useful adenoviral vectors are described in U.S. Patent Number 5,707,618 to Armentano ("Adenovirus vectors for gene therapy"), hereby incorporated by reference. For adenoviral vectors, see also Antinozzi, P.A. et al. (1999) Annu. Rev. Nutr. 19:511-544; and Verma, I.M. and N. Somia (1997) Nature 18:389:239-242, both incorporated by reference herein.

In another alternative, a herpes-based, gene therapy delivery system is used to deliver polynucleotides encoding ORP to target cells which have one or more genetic abnormalities with respect to the expression of ORP. The use of herpes simplex virus (HSV)-based vectors may be especially valuable for introducing ORP to cells of the central nervous system, for which HSV has a tropism. The construction and packaging of herpes-based vectors are well known to those with ordinary skill in the art. A replication-competent herpes simplex virus (HSV) type 1-based vector has been used to deliver a reporter gene to the eyes of primates (Liu, X. et al. (1999) Exp. Eye Res. 169:385-395). The construction of a HSV-1 virus vector has also been disclosed in detail in U.S. Patent Number 5,804,413 to DeLuca ("Herpes simplex virus strains for gene transfer"), which is hereby incorporated by reference. U.S. Patent Number 5,804,413 teaches the use of recombinant HSV d92 which consists of a genome containing at least one exogenous gene to be transferred to a cell under the control of the appropriate promoter for purposes including human gene therapy. Also taught by this patent are the

construction and use of recombinant HSV strains deleted for ICP4, ICP27 and ICP22. For HSV vectors, see also Goins, W.F. et al. (1999) *J. Virol.* 73:519-532 and Xu, H. et al. (1994) *Dev. Biol.* 163:152-161, hereby incorporated by reference. The manipulation of cloned herpesvirus sequences, the generation of recombinant virus following the transfection of multiple plasmids containing different segments of the large herpesvirus genomes, the growth and propagation of herpesvirus, and the infection of cells with herpesvirus are techniques well known to those of ordinary skill in the art.

In another alternative, an alphavirus (positive, single-stranded RNA virus) vector is used to deliver polynucleotides encoding ORP to target cells. The biology of the prototypic alphavirus, Semliki Forest Virus (SFV), has been studied extensively and gene transfer vectors have been based on the SFV genome (Garoff, H. and K.-J. Li (1998) *Curr. Opin. Biotechnol.* 9:464-469). During alphavirus RNA replication, a subgenomic RNA is generated that normally encodes the viral capsid proteins. This subgenomic RNA replicates to higher levels than the full-length genomic RNA, resulting in the overproduction of capsid proteins relative to the viral proteins with enzymatic activity (e.g., protease and polymerase). Similarly, inserting the coding sequence for ORP into the alphavirus genome in place of the capsid-coding region results in the production of a large number of ORP-coding RNAs and the synthesis of high levels of ORP in vector transduced cells. While alphavirus infection is typically associated with cell lysis within a few days, the ability to establish a persistent infection in hamster normal kidney cells (BHK-21) with a variant of Sindbis virus (SIN) indicates that the lytic replication of alphaviruses can be altered to suit the needs of the gene therapy application (Dryga, S.A. et al. (1997) *Virology* 228:74-83). The wide host range of alphaviruses will allow the introduction of ORP into a variety of cell types. The specific transduction of a subset of cells in a population may require the sorting of cells prior to transduction. The methods of manipulating infectious cDNA clones of alphaviruses, performing alphavirus cDNA and RNA transfections, and performing alphavirus infections, are well known to those with ordinary skill in the art.

Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, may also be employed to inhibit gene expression. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes; enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme

molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding ORP.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by *in vitro* and *in vivo* transcription of DNA sequences encoding ORP. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

An additional embodiment of the invention encompasses a method for screening for a compound which is effective in altering expression of a polynucleotide encoding ORP. Compounds which may be effective in altering expression of a specific polynucleotide may include, but are not limited to, oligonucleotides, antisense oligonucleotides, triple helix-forming oligonucleotides, transcription factors and other polypeptide transcriptional regulators, and non-macromolecular chemical entities which are capable of interacting with specific polynucleotide sequences. Effective compounds may alter polynucleotide expression by acting as either inhibitors or promoters of polynucleotide expression. Thus, in the treatment of disorders associated with increased ORP expression or activity, a compound which specifically inhibits expression of the polynucleotide encoding ORP may be therapeutically useful, and in the treatment of disorders associated with decreased ORP expression or activity, a compound which specifically promotes expression of the

polynucleotide encoding ORP may be therapeutically useful.

At least one, and up to a plurality, of test compounds may be screened for effectiveness in altering expression of a specific polynucleotide. A test compound may be obtained by any method commonly known in the art, including chemical modification of a compound known to be effective in altering polynucleotide expression; selection from an existing, commercially-available or proprietary library of naturally-occurring or non-natural chemical compounds; rational design of a compound based on chemical and/or structural properties of the target polynucleotide; and selection from a library of chemical compounds created combinatorially or randomly. A sample comprising a polynucleotide encoding ORP is exposed to at least one test compound thus obtained. The sample may comprise, for example, an intact or permeabilized cell, or an in vitro cell-free or reconstituted biochemical system. Alterations in the expression of a polynucleotide encoding ORP are assayed by any method commonly known in the art. Typically, the expression of a specific nucleotide is detected by hybridization with a probe having a nucleotide sequence complementary to the sequence of the polynucleotide encoding ORP. The amount of hybridization may be quantified, thus forming the basis for a comparison of the expression of the polynucleotide both with and without exposure to one or more test compounds. Detection of a change in the expression of a polynucleotide exposed to a test compound indicates that the test compound is effective in altering the expression of the polynucleotide. A screen for a compound effective in altering expression of a specific polynucleotide can be carried out, for example, using a Schizosaccharomyces pombe gene expression system (Atkins, D. et al. (1999) U.S. Patent No. 5,932,435; Arndt, G.M. et al. (2000) Nucleic Acids Res. 28:E15) or a human cell line such as HeLa cell (Clarke, M.L. et al. (2000) Biochem. Biophys. Res. Commun. 268:8-13). A particular embodiment of the present invention involves screening a combinatorial library of oligonucleotides (such as deoxyribonucleotides, ribonucleotides, peptide nucleic acids, and modified oligonucleotides) for antisense activity against a specific polynucleotide sequence (Bruce, T.W. et al. (1997) U.S. Patent No. 5,686,242; Bruce, T.W. et al. (2000) U.S. Patent No. 6,022,691).

Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nat. Biotechnol. 15:462-466.)

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as humans, dogs, cats, cows, horses, rabbits, and monkeys.

An additional embodiment of the invention relates to the administration of a composition which

generally comprises an active ingredient formulated with a pharmaceutically acceptable excipient. Excipients may include, for example, sugars, starches, celluloses, gums, and proteins. Various formulations are commonly known and are thoroughly discussed in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA). Such compositions may consist of ORP, antibodies to ORP, and mimetics, agonists, antagonists, or inhibitors of ORP.

The compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, pulmonary, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

Compositions for pulmonary administration may be prepared in liquid or dry powder form. These compositions are generally aerosolized immediately prior to inhalation by the patient. In the case of small molecules (e.g. traditional low molecular weight organic drugs), aerosol delivery of fast-acting formulations is well-known in the art. In the case of macromolecules (e.g. larger peptides and proteins), recent developments in the field of pulmonary delivery via the alveolar region of the lung have enabled the practical delivery of drugs such as insulin to blood circulation (see, e.g., Patton, J.S. et al., U.S. Patent No. 5,997,848). Pulmonary delivery has the advantage of administration without needle injection, and obviates the need for potentially toxic penetration enhancers.

Compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

Specialized forms of compositions may be prepared for direct intracellular delivery of macromolecules comprising ORP or fragments thereof. For example, liposome preparations containing a cell-impermeable macromolecule may promote cell fusion and intracellular delivery of the macromolecule. Alternatively, ORP or a fragment thereof may be joined to a short cationic N-terminal portion from the HIV Tat-1 protein. Fusion proteins thus generated have been found to transduce into the cells of all tissues, including the brain, in a mouse model system (Schwarze, S.R. et al. (1999) Science 285:1569-1572).

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells, or in animal models such as mice, rats, rabbits, dogs, monkeys, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example ORP or fragments thereof, antibodies of ORP, and agonists, antagonists or inhibitors of ORP, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by

standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED_{50} (the dose therapeutically effective in 50% of the population) or LD_{50} (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the LD_{50}/ED_{50} ratio. Compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED_{50} with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

- 10 The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy.
- 15 Long-acting compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about 0.1 g to 100,000 g, up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art.

- 20 Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

DIAGNOSTICS

- In another embodiment, antibodies which specifically bind ORP may be used for the diagnosis of disorders characterized by expression of ORP, or in assays to monitor patients being treated with ORP or agonists, antagonists, or inhibitors of ORP. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for ORP include methods which utilize the antibody and a label to detect ORP in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

- A variety of protocols for measuring ORP, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of ORP expression. Normal or standard values for ORP expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, for example, human subjects, with antibody to ORP under conditions

suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of ORP expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

5 In another embodiment of the invention, the polynucleotides encoding ORP may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantify gene expression in biopsied tissues in which expression of ORP may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of
10 ORP, and to monitor regulation of ORP levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding ORP or closely related molecules may be used to identify nucleic acid sequences which encode ORP. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a
15 conserved motif, and the stringency of the hybridization or amplification will determine whether the probe identifies only naturally occurring sequences encoding ORP, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and may have at least 50% sequence identity to any of the ORP encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:28-54 or from
20 genomic sequences including promoters, enhancers, and introns of the ORP gene.

Means for producing specific hybridization probes for DNAs encoding ORP include the cloning of polynucleotide sequences encoding ORP or ORP derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes *in vitro* by means of the addition of the appropriate RNA polymerases and the
25 appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as ^{32}P or ^{35}S , or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding ORP may be used for the diagnosis of disorders associated with expression of ORP. Examples of such disorders include, but are not limited to, a cell proliferative
30 disorder, such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease, myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia; a cancer, such as adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart,
35 kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen,

testis, thymus, thyroid, and uterus; an endocrine disorder, such as disorders of the hypothalamus and pituitary resulting from lesions such as primary brain tumors, adenomas, infarction associated with pregnancy, hypophysectomy, aneurysms, vascular malformations, thrombosis, infections, immunological disorders, and complications due to head trauma, disorders associated with

5 hypopituitarism including hypogonadism, Sheehan syndrome, diabetes insipidus, Kallman's disease, Hand-Schuller-Christian disease, Letterer-Siwe disease, sarcoidosis, empty sella syndrome, and dwarfism, disorders associated with hyperpituitarism including acromegaly, gigantism, and syndrome of inappropriate antidiuretic hormone (ADH) secretion (SIADH) often caused by benign adenoma, disorders associated with hypothyroidism including goiter, myxedema, acute thyroiditis associated

10 with bacterial infection, subacute thyroiditis associated with viral infection, autoimmune thyroiditis (Hashimoto's disease), and cretinism, disorders associated with hyperthyroidism including thyrotoxicosis and its various forms, Grave's disease, pretibial myxedema, toxic multinodular goiter, thyroid carcinoma, and Plummer's disease, disorders associated with hyperparathyroidism including Conn disease (chronic hypercalcemia), pancreatic disorders such as Type I or Type II diabetes mellitus

15 and associated complications, disorders associated with the adrenals such as hyperplasia, carcinoma, or adenoma of the adrenal cortex, hypertension associated with alkalosis, amyloidosis, hypokalemia, Cushing's disease, Liddle's syndrome, and Arnold-Healy-Gordon syndrome, pheochromocytoma tumors, and Addison's disease, disorders associated with gonadal steroid hormones such as: in women, abnormal prolactin production, infertility, endometriosis, perturbations of the menstrual

20 cycle, polycystic ovarian disease, hyperprolactinemia, isolated gonadotropin deficiency, amenorrhea, galactorrhea, hermaphroditism, hirsutism and virilization, breast cancer, and, in post-menopausal women, osteoporosis, and, in men, Leydig cell deficiency, male climacteric phase, and germinal cell aplasia, hypergonadal disorders associated with Leydig cell tumors, androgen resistance associated with absence of androgen receptors, syndrome of 5 α -reductase, and gynecomastia; a metabolic

25 disorder, such as Addison's disease, cystic fibrosis, diabetes, fatty hepatocirrhosis, galactosemia, goiter, hyperadrenalism, hypoadrenalism, hyperparathyroidism, hypoparathyroidism, hypercholesterolemia, hyperthyroidism, hypothyroidism hyperlipidemia, hyperlipemia, lipid myopathies, obesity, lipodystrophies, and phenylketonuria, congenital adrenal hyperplasia, pseudovitamin D-deficiency rickets, cerebrotendinous xanthomatosis, and coumarin resistance; a

30 reproductive disorder, such as disorders of prolactin production, infertility, including tubal disease, ovulatory defects, and endometriosis, disruptions of the estrous cycle, disruptions of the menstrual cycle, polycystic ovary syndrome, ovarian hyperstimulation syndrome, endometrial and ovarian tumors, uterine fibroids, autoimmune disorders, ectopic pregnancies, and teratogenesis; cancer of the breast, fibrocystic breast disease, and galactorrhea, disruptions of spermatogenesis, abnormal sperm

35 physiology, cancer of the testis, cancer of the prostate, benign prostatic hyperplasia, prostatitis,

Peyronie's disease, impotence, carcinoma of the male breast, and gynecomastia; a neurological disorder, such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathisia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, and Tourette's disorder; an autoimmune/inflammatory disorder, such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; and a viral disorder, such as viral infections, e.g., those caused by adenoviruses (acute respiratory disease, pneumonia), arenaviruses (lymphocytic choriomeningitis), bunyaviruses (Hantavirus), coronaviruses (pneumonia, chronic bronchitis), hepadnaviruses (hepatitis), herpesviruses (herpes simplex virus, varicella-zoster virus, Epstein-Barr virus, cytomegalovirus), flaviviruses (yellow fever), orthomyxoviruses (influenza),

papillomaviruses (cancer), paramyxoviruses (measles, mumps), picornoviruses (rhinovirus, poliovirus, coxsackie-virus), polyomaviruses (BK virus, JC virus), poxviruses (smallpox), reovirus (Colorado tick fever), retroviruses (human immunodeficiency virus, human T lymphotropic virus), rhabdoviruses (rabies), rotaviruses (gastroenteritis), and togaviruses (encephalitis, rubella). The polynucleotide sequences encoding ORP may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered ORP expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding ORP may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding ORP may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantified and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding ORP in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of ORP, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding ORP, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of an abnormal amount of transcript (either under- or overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual

clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding ORP may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced in vitro. Oligomers will preferably contain a fragment of a polynucleotide encoding ORP, or a fragment of a polynucleotide complementary to the polynucleotide encoding ORP, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantification of closely related DNA or RNA sequences.

In a particular aspect, oligonucleotide primers derived from the polynucleotide sequences encoding ORP may be used to detect single nucleotide polymorphisms (SNPs). SNPs are substitutions, insertions and deletions that are a frequent cause of inherited or acquired genetic disease in humans. Methods of SNP detection include, but are not limited to, single-stranded conformation polymorphism (SSCP) and fluorescent SSCP (fSSCP) methods. In SSCP, oligonucleotide primers derived from the polynucleotide sequences encoding ORP are used to amplify DNA using the polymerase chain reaction (PCR). The DNA may be derived, for example, from diseased or normal tissue, biopsy samples, bodily fluids, and the like. SNPs in the DNA cause differences in the secondary and tertiary structures of PCR products in single-stranded form, and these differences are detectable using gel electrophoresis in non-denaturing gels. In fSSCP, the oligonucleotide primers are fluorescently labeled, which allows detection of the amplimers in high-throughput equipment such as DNA sequencing machines. Additionally, sequence database analysis methods, termed *in silico* SNP (isSNP), are capable of identifying polymorphisms by comparing the sequence of individual overlapping DNA fragments which assemble into a common consensus sequence. These computer-based methods filter out sequence variations due to laboratory preparation of DNA and sequencing errors using statistical models and automated analyses of DNA sequence chromatograms. In the alternative, SNPs may be detected and characterized by mass spectrometry using, for example, the high throughput MASSARRAY system (Sequenom, Inc., San Diego CA).

Methods which may also be used to quantify the expression of ORP include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) J. Immunol. Methods 159:235-244; Duplaa, C. et al. (1993) Anal. Biochem. 212:229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in a high-throughput format where the oligomer or polynucleotide of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as elements on a microarray. The microarray can be used in transcript imaging techniques which monitor the relative expression levels of large numbers of genes simultaneously as described in Seilhamer, J.J. et al., "Comparative Gene Transcript Analysis," U.S. Patent No. 5,840,484, incorporated herein by reference. The microarray may also be used to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, to monitor progression/regression of disease as a function of gene expression, and to develop and monitor the activities of therapeutic agents in the treatment of disease. In particular, this information may be used to develop a pharmacogenomic profile of a patient in order to select the most appropriate and effective treatment regimen for that patient. For example, therapeutic agents which are highly effective and display the fewest side effects may be selected for a patient based on his/her pharmacogenomic profile.

In another embodiment, antibodies specific for ORP, or ORP or fragments thereof may be used as elements on a microarray. The microarray may be used to monitor or measure protein-protein interactions, drug-target interactions, and gene expression profiles, as described above.

A particular embodiment relates to the use of the polynucleotides of the present invention to generate a transcript image of a tissue or cell type. A transcript image represents the global pattern of gene expression by a particular tissue or cell type. Global gene expression patterns are analyzed by quantifying the number of expressed genes and their relative abundance under given conditions and at a given time. (See Seilhamer et al., "Comparative Gene Transcript Analysis," U.S. Patent Number 5,840,484, expressly incorporated by reference herein.) Thus a transcript image may be generated by hybridizing the polynucleotides of the present invention or their complements to the totality of transcripts or reverse transcripts of a particular tissue or cell type. In one embodiment, the hybridization takes place in high-throughput format, wherein the polynucleotides of the present invention or their complements comprise a subset of a plurality of elements on a microarray. The resultant transcript image would provide a profile of gene activity.

Transcript images may be generated using transcripts isolated from tissues, cell lines, biopsies, or other biological samples. The transcript image may thus reflect gene expression in vivo, as in the case of a tissue or biopsy sample, or in vitro, as in the case of a cell line.

Transcript images which profile the expression of the polynucleotides of the present invention may also be used in conjunction with in vitro model systems and preclinical evaluation of pharmaceuticals, as well as toxicological testing of industrial and naturally-occurring environmental compounds. All compounds induce characteristic gene expression patterns, frequently termed molecular fingerprints or toxicant signatures, which are indicative of mechanisms of action and toxicity (Nuwaysir, E.F. et al. (1999) Mol. Carcinog. 24:153-159; Steiner, S. and N.L. Anderson (2000)

Toxicol. Lett. 112-113:467-471, expressly incorporated by reference herein). If a test compound has a signature similar to that of a compound with known toxicity, it is likely to share those toxic properties. These fingerprints or signatures are most useful and refined when they contain expression information from a large number of genes and gene families. Ideally, a genome-wide measurement of expression provides the highest quality signature. Even genes whose expression is not altered by any tested compounds are important as well, as the levels of expression of these genes are used to normalize the rest of the expression data. The normalization procedure is useful for comparison of expression data after treatment with different compounds. While the assignment of gene function to elements of a toxicant signature aids in interpretation of toxicity mechanisms, knowledge of gene function is not necessary for the statistical matching of signatures which leads to prediction of toxicity. (See, for example, Press Release 00-02 from the National Institute of Environmental Health Sciences, released February 29, 2000, available at <http://www.niehs.nih.gov/oc/news/toxchip.htm>.) Therefore, it is important and desirable in toxicological screening using toxicant signatures to include all expressed gene sequences.

In one embodiment, the toxicity of a test compound is assessed by treating a biological sample containing nucleic acids with the test compound. Nucleic acids that are expressed in the treated biological sample are hybridized with one or more probes specific to the polynucleotides of the present invention, so that transcript levels corresponding to the polynucleotides of the present invention may be quantified. The transcript levels in the treated biological sample are compared with levels in an untreated biological sample. Differences in the transcript levels between the two samples are indicative of a toxic response caused by the test compound in the treated sample.

Another particular embodiment relates to the use of the polypeptide sequences of the present invention to analyze the proteome of a tissue or cell type. The term proteome refers to the global pattern of protein expression in a particular tissue or cell type. Each protein component of a proteome can be subjected individually to further analysis. Proteome expression patterns, or profiles, are analyzed by quantifying the number of expressed proteins and their relative abundance under given conditions and at a given time. A profile of a cell's proteome may thus be generated by separating and analyzing the polypeptides of a particular tissue or cell type. In one embodiment, the separation is achieved using two-dimensional gel electrophoresis, in which proteins from a sample are separated by isoelectric focusing in the first dimension, and then according to molecular weight by sodium dodecyl sulfate slab gel electrophoresis in the second dimension (Steiner and Anderson, *supra*). The proteins are visualized in the gel as discrete and uniquely positioned spots, typically by staining the gel with an agent such as Coomassie Blue or silver or fluorescent stains. The optical density of each protein spot is generally proportional to the level of the protein in the sample. The optical densities of equivalently positioned protein spots from different samples, for example, from biological samples either treated or

untreated with a test compound or therapeutic agent, are compared to identify any changes in protein spot density related to the treatment. The proteins in the spots are partially sequenced using, for example, standard methods employing chemical or enzymatic cleavage followed by mass spectrometry. The identity of the protein in a spot may be determined by comparing its partial sequence, preferably of at least 5 contiguous amino acid residues, to the polypeptide sequences of the present invention. In some cases, further sequence data may be obtained for definitive protein identification.

A proteomic profile may also be generated using antibodies specific for ORP to quantify the levels of ORP expression. In one embodiment, the antibodies are used as elements on a microarray, and protein expression levels are quantified by exposing the microarray to the sample and detecting the levels of protein bound to each array element (Lueking, A. et al. (1999) *Anal. Biochem.* 270:103-111; Mendoz, L.G. et al. (1999) *Biotechniques* 27:778-788). Detection may be performed by a variety of methods known in the art, for example, by reacting the proteins in the sample with a thiol- or amino-reactive fluorescent compound and detecting the amount of fluorescence bound at each array element.

Toxicant signatures at the proteome level are also useful for toxicological screening, and should be analyzed in parallel with toxicant signatures at the transcript level. There is a poor correlation between transcript and protein abundances for some proteins in some tissues (Anderson, N.L. and J. Seilhamer (1997) *Electrophoresis* 18:533-537), so proteome toxicant signatures may be useful in the analysis of compounds which do not significantly affect the transcript image, but which alter the proteomic profile. In addition, the analysis of transcripts in body fluids is difficult, due to rapid degradation of mRNA, so proteomic profiling may be more reliable and informative in such cases.

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins that are expressed in the treated biological sample are separated so that the amount of each protein can be quantified. The amount of each protein is compared to the amount of the corresponding protein in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample. Individual proteins are identified by sequencing the amino acid residues of the individual proteins and comparing these partial sequences to the polypeptides of the present invention.

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins from the biological sample are incubated with antibodies specific to the polypeptides of the present invention. The amount of protein recognized by the antibodies is quantified. The amount of protein in the treated biological sample is compared with the amount in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) *Proc. Natl. Acad. Sci.*

USA 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.) Various types of microarrays are well known and thoroughly described in DNA Microarrays: A Practical Approach, M. Schena, ed. (1999) Oxford University Press, London, hereby expressly incorporated by reference.

In another embodiment of the invention, nucleic acid sequences encoding ORP may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. Either coding or noncoding sequences may be used, and in some instances, noncoding sequences may be preferable over coding sequences. For example, conservation of a coding sequence among members of a multi-gene family may potentially cause undesired cross hybridization during chromosomal mapping. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.) Once mapped, the nucleic acid sequences of the invention may be used to develop genetic linkage maps, for example, which correlate the inheritance of a disease state with the inheritance of a particular chromosome region or restriction fragment length polymorphism (RFLP). (See, e.g., Lander, E.S. and D. Botstein (1986) Proc. Natl. Acad. Sci. USA 83:7353-7357.)

Fluorescent in situ hybridization (FISH) may be correlated with other physical and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, supra, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) World Wide Web site. Correlation between the location of the gene encoding ORP on a physical map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder and thus may further positional cloning efforts.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the exact chromosomal locus is not known. This information is valuable to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the gene or genes responsible for a disease or syndrome have been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) Nature 336:577-580.) The nucleotide sequence of the instant invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc.,

among normal, carrier, or affected individuals.

In another embodiment of the invention, ORP, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between ORP and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with ORP, or fragments thereof, and washed. Bound ORP is then detected by methods well known in the art. Purified ORP can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding ORP specifically compete with a test compound for binding ORP. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with ORP.

In additional embodiments, the nucleotide sequences which encode ORP may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications, and publications mentioned above and below, in particular U.S. Ser. No. 60/172,367, are hereby expressly incorporated by reference.

EXAMPLES

I. Construction of cDNA Libraries

RNA was purchased from Clontech or isolated from tissues described in Table 4. Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol

or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A+) RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN, Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERScript plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, *supra*, units 5.1-6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), PSPORT1 plasmid (Life Technologies), pcDNA2.1 plasmid (Invitrogen, Carlsbad CA), or pINCY plasmid (Incyte Genomics, Palo Alto CA). Recombinant plasmids were transformed into competent *E. coli* cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5 α , DH10B, or ElectroMAX DH10B from Life Technologies.

II. Isolation of cDNA Clones

Plasmids obtained as described in Example I were recovered from host cells by *in vivo* excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96 plasmid purification kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4 C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a FLUOROSKAN II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

III. Sequencing and Analysis

Incyte cDNA recovered in plasmids as described in Example II were sequenced as follows. Sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (Applied Biosystems) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing kits such as the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 sequencing system (Applied Biosystems) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA sequences were identified using standard methods (reviewed in Ausubel, 1997, *supra*, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example VI.

The polynucleotide sequences derived from cDNA sequencing were assembled and analyzed using a combination of software programs which utilize algorithms well known to those skilled in the art. Table 5 summarizes the tools, programs, and algorithms used and provides applicable descriptions, references, and threshold parameters. The first column of Table 5 shows the tools, programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score, the greater the homology between two sequences). Sequences were analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments were generated using the default parameters specified by the clustal algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.

The polynucleotide sequences were validated by removing vector, linker, and polyA sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programming, and dinucleotide nearest neighbor analysis. The sequences were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS, PRINTS, DOMO, PRODOM, and PFAM to acquire annotation using programs based on BLAST, FASTA, and BLIMPS. The sequences were assembled into full length polynucleotide sequences using programs based on Phred, Phrap, and Consed, and were screened

for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length amino acid sequences, and these full length sequences were subsequently analyzed by querying against databases such as the GenBank databases (described above), SwissProt, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, and Hidden Markov Model (HMM)-based protein family databases such as PFAM. HMM is a probabilistic approach which analyzes consensus primary structures of gene families. (See, e.g., Eddy, S.R. (1996) *Curr. Opin. Struct. Biol.* 6:361-365.)

The programs described above for the assembly and analysis of full length polynucleotide and amino acid sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:28-54. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies were described in The Invention section above.

IV. Analysis of Polynucleotide Expression

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, *supra*, ch. 7; Ausubel, 1995, *supra*, ch. 4 and 16.)

Analogous computer techniques applying BLAST were used to search for identical or related molecules in cDNA databases such as GenBank or LIFESEQ (Incyte Genomics). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$\frac{\text{BLAST Score} \times \text{Percent Identity}}{5 \times \text{minimum} \{ \text{length}(\text{Seq. 1}), \text{length}(\text{Seq. 2}) \}}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. The product score is a normalized value between 0 and 100, and is calculated as follows: the BLAST score is multiplied by the percent nucleotide identity and the product is divided by (5 times the length of the shorter of the two sequences). The BLAST score is calculated by assigning a score of +5 for every base that matches in a high-scoring segment pair (HSP), and -4 for every mismatch. Two sequences may share more than one HSP (separated by gaps). If there is more than one HSP, then the pair with the highest BLAST score is used to calculate the product score. The product score represents a balance between fractional overlap and quality in a BLAST alignment. For example, a product score of 100 is produced only for 100% identity over the entire length of the shorter of the two sequences being compared. A product score of 70 is produced either by 100% identity and 70% overlap at one end, or by 88% identity and 100% overlap at the other. A product score of 50 is

produced either by 100% identity and 50% overlap at one end, or 79% identity and 100% overlap.

The results of northern analyses are reported as a percentage distribution of libraries in which the transcript encoding ORP occurred. Analysis involved the categorization of cDNA libraries by organ/tissue and disease. The organ/tissue categories included cardiovascular, dermatologic, developmental, endocrine, gastrointestinal, hematopoietic/immune, musculoskeletal, nervous, reproductive, and urologic. The disease/condition categories included cancer, inflammation, trauma, cell proliferation, neurological, and pooled. For each category, the number of libraries expressing the sequence of interest was counted and divided by the total number of libraries across all categories. Percentage values of tissue-specific and disease- or condition-specific expression are reported in Table 3.

V. Chromosomal Mapping of ORP Encoding Polynucleotides

The cDNA sequences which were used to assemble SEQ ID NO:28-54 were compared with sequences from the Incyte LIFESEQ database and public domain databases using BLAST and other implementations of the Smith-Waterman algorithm. Sequences from these databases that matched SEQ ID NO:28-54 were assembled into clusters of contiguous and overlapping sequences using assembly algorithms such as Phrap (Table 5). Radiation hybrid and genetic mapping data available from public resources such as the Stanford Human Genome Center (SHGC), Whitehead Institute for Genome Research (WIGR), and Généthon were used to determine if any of the clustered sequences had been previously mapped. Inclusion of a mapped sequence in a cluster resulted in the assignment of all sequences of that cluster, including its particular SEQ ID NO., to that map location.

The genetic map locations of SEQ ID NO:31, SEQ ID NO:42, SEQ ID NO:48, and SEQ ID NO:53 are described in The Invention as ranges, or intervals, of human chromosomes. More than one map location is reported for SEQ ID NO:48, indicating that previously mapped sequences having similarity, but not complete identity, to SEQ ID NO:48 were assembled into their respective clusters. The map position of an interval, in centiMorgans, is measured relative to the terminus of the chromosome's p-arm. (The centiMorgan (cM) is a unit of measurement based on recombination frequencies between chromosomal markers. On average, 1 cM is roughly equivalent to 1 megabase (Mb) of DNA in humans, although this can vary widely due to hot and cold spots of recombination.) The cM distances are based on genetic markers mapped by Généthon which provide boundaries for radiation hybrid markers whose sequences were included in each of the clusters. Human genome maps and other resources available to the public, such as the NCBI "GeneMap'99" World Wide Web site (<http://www.ncbi.nlm.nih.gov/genemap/>), can be employed to determine if previously identified disease genes map within or in proximity to the intervals indicated above.

VI. Extension of ORP Encoding Polynucleotides

The full length nucleic acid sequences of SEQ ID NO:28-54 were produced by extension of an

appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer, to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68 °C to about 72 °C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg^{2+} , $(NH_4)_2SO_4$, and β -mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94 °C, 3 min; Step 2: 94 °C, 15 sec; Step 3: 60 °C, 1 min; Step 4: 68 °C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68 °C, 5 min; Step 7: storage at 4 °C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94 °C, 3 min; Step 2: 94 °C, 15 sec; Step 3: 57 °C, 1 min; Step 4: 68 °C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68 °C, 5 min; Step 7: storage at 4 °C.

The concentration of DNA in each well was determined by dispensing 100 μ l PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5 μ l of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5 μ l to 10 μ l aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose mini-gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent *E. coli* cells. Transformed cells were selected on antibiotic-containing media, and individual colonies were picked and cultured overnight at 37 °C in 384-well plates in LB/2x

carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94 C, 3 min; Step 2: 94 C, 15 sec; Step 3: 60 C, 1 min; Step 4: 72 C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72 C, 5 min; Step 7: storage at 4 C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethylsulfoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems).

In like manner, the polynucleotide sequences of SEQ ID NO:28-54 are used to obtain 5' regulatory sequences using the procedure above, along with oligonucleotides designed for such extension, and an appropriate genomic library.

VII. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:28-54 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 Ci of [γ -³²P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10⁷ counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40 C. To remove nonspecific signals, blots are sequentially washed at room temperature under conditions of up to, for example, 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. Hybridization patterns are visualized using autoradiography or an alternative imaging means and compared.

VIII. Microarrays

The linkage or synthesis of array elements upon a microarray can be achieved utilizing photolithography, piezoelectric printing (ink-jet printing, See, e.g., Baldeschweiler, supra), mechanical microspotting technologies, and derivatives thereof. The substrate in each of the aforementioned

technologies should be uniform and solid with a non-porous surface (Skena (1999), supra). Suggested substrates include silicon, silica, glass slides, glass chips, and silicon wafers. Alternatively, a procedure analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced using available methods and machines well known to those of ordinary skill in the art and may contain any appropriate number of elements. (See, e.g., Skena, M. et al. (1995) Science 270:467-470; Shalon, D. et al. (1996) Genome Res. 6:639-645; Marshall, A. and J. Hodgson (1998) Nat. Biotechnol. 16:27-31.)

Full length cDNAs, Expressed Sequence Tags (ESTs), or fragments or oligomers thereof may comprise the elements of the microarray. Fragments or oligomers suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). The array elements are hybridized with polynucleotides in a biological sample. The polynucleotides in the biological sample are conjugated to a fluorescent label or other molecular tag for ease of detection. After hybridization, nonhybridized nucleotides from the biological sample are removed, and a fluorescence scanner is used to detect hybridization at each array element. Alternatively, laser desorption and mass spectrometry may be used for detection of hybridization. The degree of complementarity and the relative abundance of each polynucleotide which hybridizes to an element on the microarray may be assessed. In one embodiment, microarray preparation and usage is described in detail below.

Tissue or Cell Sample Preparation

Total RNA is isolated from tissue samples using the guanidinium thiocyanate method and poly(A)⁺ RNA is purified using the oligo-(dT) cellulose method. Each poly(A)⁺ RNA sample is reverse transcribed using MMLV reverse-transcriptase, 0.05 pg/μl oligo-(dT) primer (21mer), 1X first strand buffer, 0.03 units/μl RNase inhibitor, 500 μM dATP, 500 μM dGTP, 500 μM dTTP, 40 μM dCTP, 40 μM dCTP-Cy3 (BDS) or dCTP-Cy5 (Amersham Pharmacia Biotech). The reverse transcription reaction is performed in a 25 ml volume containing 200 ng poly(A)⁺ RNA with GEMBRIGHT kits (Incyte). Specific control poly(A)⁺ RNAs are synthesized by in vitro transcription from non-coding yeast genomic DNA. After incubation at 37 °C for 2 hr, each reaction sample (one with Cy3 and another with Cy5 labeling) is treated with 2.5 ml of 0.5M sodium hydroxide and incubated for 20 minutes at 85 °C to stop the reaction and degrade the RNA. Samples are purified using two successive CHROMA SPIN 30 gel filtration spin columns (CLONTECH Laboratories, Inc. (CLONTECH), Palo Alto CA) and after combining, both reaction samples are ethanol precipitated using 1 ml of glycogen (1 mg/ml), 60 ml sodium acetate, and 300 ml of 100% ethanol. The sample is then dried to completion using a SpeedVAC (Savant Instruments Inc., Holbrook NY) and resuspended in 14 μl 5X SSC/0.2% SDS.

Microarray Preparation

Sequences of the present invention are used to generate array elements. Each array element is amplified from bacterial cells containing vectors with cloned cDNA inserts. PCR amplification uses primers complementary to the vector sequences flanking the cDNA insert. Array elements are amplified in thirty cycles of PCR from an initial quantity of 1-2 ng to a final quantity greater than 5 μ g. Amplified array elements are then purified using SEPHACRYL-400 (Amersham Pharmacia Biotech).

Purified array elements are immobilized on polymer-coated glass slides. Glass microscope slides (Corning) are cleaned by ultrasound in 0.1% SDS and acetone, with extensive distilled water washes between and after treatments. Glass slides are etched in 4% hydrofluoric acid (VWR Scientific Products Corporation (VWR), West Chester PA), washed extensively in distilled water, and coated with 0.05% aminopropyl silane (Sigma) in 95% ethanol. Coated slides are cured in a 110°C oven.

Array elements are applied to the coated glass substrate using a procedure described in US Patent No. 5,807,522, incorporated herein by reference. 1 μ l of the array element DNA, at an average concentration of 100 ng/ μ l, is loaded into the open capillary printing element by a high-speed robotic apparatus. The apparatus then deposits about 5 nl of array element sample per slide.

Microarrays are UV-crosslinked using a STRATALINKER UV-crosslinker (Stratagene). Microarrays are washed at room temperature once in 0.2% SDS and three times in distilled water. Non-specific binding sites are blocked by incubation of microarrays in 0.2% casein in phosphate buffered saline (PBS) (Tropix, Inc., Bedford MA) for 30 minutes at 60 °C followed by washes in 0.2% SDS and distilled water as before.

Hybridization

Hybridization reactions contain 9 μ l of sample mixture consisting of 0.2 μ g each of Cy3 and Cy5 labeled cDNA synthesis products in 5X SSC, 0.2% SDS hybridization buffer. The sample mixture is heated to 65 °C for 5 minutes and is aliquoted onto the microarray surface and covered with an 1.8 cm² coverslip. The arrays are transferred to a waterproof chamber having a cavity just slightly larger than a microscope slide. The chamber is kept at 100% humidity internally by the addition of 140 μ l of 5X SSC in a corner of the chamber. The chamber containing the arrays is incubated for about 6.5 hours at 60 °C. The arrays are washed for 10 min at 45 °C in a first wash buffer (1X SSC, 0.1% SDS), three times for 10 minutes each at 45 °C in a second wash buffer (0.1X SSC), and dried.

Detection

Reporter-labeled hybridization complexes are detected with a microscope equipped with an Innova 70 mixed gas 10 W laser (Coherent, Inc., Santa Clara CA) capable of generating spectral lines at 488 nm for excitation of Cy3 and at 632 nm for excitation of Cy5. The excitation laser light is

focused on the array using a 20X microscope objective (Nikon, Inc., Melville NY). The slide containing the array is placed on a computer-controlled X-Y stage on the microscope and raster-scanned past the objective. The 1.8 cm x 1.8 cm array used in the present example is scanned with a resolution of 20 micrometers.

5 In two separate scans, a mixed gas multiline laser excites the two fluorophores sequentially. Emitted light is split, based on wavelength, into two photomultiplier tube detectors (PMT R1477, Hamamatsu Photonics Systems, Bridgewater NJ) corresponding to the two fluorophores. Appropriate filters positioned between the array and the photomultiplier tubes are used to filter the signals. The emission maxima of the fluorophores used are 565 nm for Cy3 and 650 nm for Cy5. Each array is
10 typically scanned twice, one scan per fluorophore using the appropriate filters at the laser source, although the apparatus is capable of recording the spectra from both fluorophores simultaneously.

The sensitivity of the scans is typically calibrated using the signal intensity generated by a cDNA control species added to the sample mixture at a known concentration. A specific location on the array contains a complementary DNA sequence, allowing the intensity of the signal at that
15 location to be correlated with a weight ratio of hybridizing species of 1:100,000. When two samples from different sources (e.g., representing test and control cells), each labeled with a different fluorophore, are hybridized to a single array for the purpose of identifying genes that are differentially expressed, the calibration is done by labeling samples of the calibrating cDNA with the two fluorophores and adding identical amounts of each to the hybridization mixture.

20 The output of the photomultiplier tube is digitized using a 12-bit RTI-835H analog-to-digital (A/D) conversion board (Analog Devices, Inc., Norwood MA) installed in an IBM-compatible PC computer. The digitized data are displayed as an image where the signal intensity is mapped using a linear 20-color transformation to a pseudocolor scale ranging from blue (low signal) to red (high signal). The data is also analyzed quantitatively. Where two different fluorophores are excited and
25 measured simultaneously, the data are first corrected for optical crosstalk (due to overlapping emission spectra) between the fluorophores using each fluorophore's emission spectrum.

A grid is superimposed over the fluorescence signal image such that the signal from each spot is centered in each element of the grid. The fluorescence signal within each element is then integrated to obtain a numerical value corresponding to the average intensity of the signal. The software used
30 for signal analysis is the GEMTOOLS gene expression analysis program (Incyte).

IX. Complementary Polynucleotides

Sequences complementary to the ORP-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring ORP. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with
35 smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO

4.06 software (National Biosciences) and the coding sequence of ORP. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the ORP-encoding transcript.

5 X. Expression of ORP

Expression and purification of ORP is achieved using bacterial or virus-based expression systems. For expression of ORP in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac* (*tac*) hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element.

Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express ORP upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of ORP in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant Autographica californica nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding ORP by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect Spodoptera frugiperda (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, ORP is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from Schistosoma japonicum, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from ORP at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, supra, ch. 10 and 16). Purified ORP obtained by these methods can be used directly in the assays shown in Examples XI and XV.

35 XI. Demonstration of ORP Activity

For purposes of example, an assay demonstrating the activity of a short-chain alcohol

dehydrogenase is described. Essentially the same method is used for other types of oxidoreductases, with suitable substitution of the substrate and electron acceptor. ORP activity is demonstrated by the oxidation of NADPH to NADP in the presence of substrate (Kunau and Dommes (1978) Eur. J.

Biochem. 91:533-544). Substrates include, but are not limited to, all-trans-retinaldehyde and *cis*-4-dienoyl-CoA. ORP is preincubated for 10 minutes at 37 °C in 60 µM potassium phosphate (pH 7.4), 125 nM NADPH, and 0.2 µM CoA (coenzyme A). The reaction is initiated by addition of the appropriate substrate (12.5 to 150 µM final concentration). The change in absorbance of the reaction at 340 nm, due to the oxidation of NADPH to NADP, is measured using a spectrophotometer at 23 °C. Units of ORP activity are expressed as µmoles of NADP formed per minute. A reaction lacking

ORP is used as a negative control.

Alternatively, ORP activity is assayed by measuring the reduction of insulin. Aliquots of ORP are preincubated at 37 °C for 20 min with 2 µl of 50 mM Hepes, pH 7.6, 100 µg/ml bovine serum albumin, and 2 mM DTT in a total volume of 70 µl. Then, 40 µl of a reaction mixture composed of 200 µl of Hepes (1 M), pH 7.6, 40 µl of EDTA (0.2 M), 40 µl of NADPH (40 mg/ml), and 500 µl of insulin (10 mg/ml) is added. The reaction is initiated with the addition of 10 µl of thioredoxin reductase from calf thymus (3.0 A412 unit), and incubation is continued for 20 min at 37 °C. The reaction rate is followed by monitoring the oxidation of NADPH at 412 nm. The oxidation of NADPH is proportional to the amount of insulin reduction.

XII. Functional Assays

ORP function is assessed by expressing the sequences encoding ORP at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include pCMV SPORT plasmid (Life Technologies) and pCR3.1 plasmid (Invitrogen), both of which contain the cytomegalovirus promoter. 5-10 µg of recombinant vector are transiently transfected into a human cell line, for example, an endothelial or hematopoietic cell line, using either liposome formulations or electroporation. 1-2 µg of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-

regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G.

5 (1994) Flow Cytometry, Oxford, New York NY.

The influence of ORP on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding ORP and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding ORP and other genes of interest can be analyzed by northern analysis or microarray techniques.

XIII. Production of ORP Specific Antibodies

15 ORP substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) *Methods Enzymol.* 182:488-495), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols.

Alternatively, the ORP amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, *supra*, ch. 11.)

Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A peptide synthesizer (Applied Biosystems) using Fmoc chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, *supra*.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide and anti-ORP activity by, for example, binding the peptide or ORP to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

30 XIV. Purification of Naturally Occurring ORP Using Specific Antibodies

Naturally occurring or recombinant ORP is substantially purified by immunoaffinity chromatography using antibodies specific for ORP. An immunoaffinity column is constructed by covalently coupling anti-ORP antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing ORP are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of ORP (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/ORP binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and ORP is collected.

XV. Identification of Molecules Which Interact with ORP

ORP, or biologically active fragments thereof, are labeled with ^{125}I Bolton-Hunter reagent. (See, e.g., Bolton A.E. and W.M. Hunter (1973) *Biochem. J.* 133:529-539.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled ORP, washed, and any wells with labeled ORP complex are assayed. Data obtained using different concentrations of ORP are used to calculate values for the number, affinity, and association of ORP with the candidate molecules.

Alternatively, molecules interacting with ORP are analyzed using the yeast two-hybrid system as described in Fields, S. and O. Song (1989, *Nature* 340:245-246), or using commercially available kits based on the two-hybrid system, such as the MATCHMAKER system (Clontech).

ORP may also be used in the PATHCALLING process (CuraGen Corp., New Haven CT) which employs the yeast two-hybrid system in a high-throughput manner to determine all interactions between the proteins encoded by two large libraries of genes (Nandabalan, K. et al. (2000) U.S. Patent No. 6,057,101).

Various modifications and variations of the described methods and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with certain embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

Table 1

Polypeptide SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
1	28	543496	OVARNOT02	195237H1 (KIDNNOT02), 849952R1 (NGANNOT01), 2053720R6 (BEPINOT01), 2763095T6 (BRSTNOT12), 4069149H1 (KIDNNOT26), 5190107T6 (OVARDT06)
2	29	907607	COLNNOT09	907607H1 (COLNNOT09), 1574381F6 (LNODNOT03), 1921146R6 (BRSTTUT01), 4362192H1 (SKIRNOT01), 4574450T6 (PROSTMT02)
3	30	1290078	BRAINOT11	1290078H1 (BRAINOT11), 1420837F1 (KIDNNOT09), 1615722F6 (BRAITUT12), 1850135F6 (LUNGFET03), 2988862H1 (CARGDIT01), 3323937F6 (PTHYNOT03), 3323937T6 (PTHYNOT03), 4131237H2 (CARGDIT01), 4132810H2 (CARGDIT01)
4	31	1302741	PLACNOT02	1223088R1 (COLNTUT02), 1302741F6 (PLACNOT02), 1302741H1 (PLACNOT02), 1804982F6 (SINTNOT13), 2946303F6 (BRAITUT23)
5	32	1541028	SINTTUT01	1480786F6 (CORPNOT02), 1482929T6 (CORPNOT02), 1541028H1 (SINTTUT01), 1541028R6 (SINTTUT01), 4801884H1 (MYEPUNT01)
6	33	1597687	BRAINOT14	752372R6 (BRAITUT01), 1597687F6 (BRAINOT14), 1597687H1 (BRAINOT14), 2536401F6 (BRAINOT18), SCFA05185V1, SCFA05069V1, SCFA05285V1
7	34	1690348	PROSTUT10	1292558T1 (PGANNOT03), 1690348H1 (PROSTUT10), 2649821F6 (THYMFET02), 3774383F6 (BRSTNOT25), 3774383T6 (BRSTNOT25)
8	35	1865603	PROSNOT19	1865603F6 (PROSNOT19), 1865603H1 (PROSNOT19), 1865603T6 (PROSNOT19), 5288406H1 (LIVRTUS02)
9	36	1976472	PANCTUT02	287078R1 (EOSIHET02), 335109T6 (EOSIHET02), 940458R1 (ADRENOT03), 1556441F1 (BLADTUT04), 1611306F6 (COLNTUT06), 1976472H1 (PANCTUT02), 2835904F6 (TLYMNOT03), 3942847F6 (SCORNOT04), 4969893H1 (KIDEUNC10), 91745222
10	37	2050821	LIVRFET02	1755049F6 (LIVRTUT01), 1985566R6 (LUNGAST01), 2050821F6 (LIVRFET02), 2050821H1 (LIVRFET02), SXBC01908V1, SXBC01618V1, 9747347, 9760098
11	38	2408443	BSTMNON02	2408443H1 (BSTMNON02)
12	39	2508668	CONUTUT01	1318620T1 (BLADNOT04), 1444080R1 (THYRNOT03), 1900006F6 (BLADTUT06), 2006550R6 (TESTNOT03), 2508668H1 (CONUTUT01)
13	40	2536830	BRAINOT18	2536830F6 (BRAINOT18), 2536830H2 (BRAINOT18), 2717227F6 (THYRNOT09), 4328825F6 (KIDNNOT32), 93804081

Table 1 (cont.)

Polypeptide SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
14	41	2645179	OVARTUT03	1272506H1 (TESTTUT02), 2467552F6 (THYRN0T08), 2467552T6 (THYRN0T08), 2645179H1 (OVARTUT03), 2729444T6 (OVARTUT05), 3035294T6 (TLYMNOT05), 3220640H1 (COLNNO03), 4606226H1 (BRSTNOT07), 5671416H1 (BONEUNT01)
15	42	2754425	THPLAZS08	804309R6 (BRAVXT05), 804309T6 (BRAVXT05), 926499T6 (BRAINOT04), 2754425H1 (THPLAZS08), 3187631H1 (THYMN004), 3244042F6 (BRAINOT19), 91231122, 91958528
16	43	2821526	ADRETUT06	288961F1 (EOSIHET02), 2821526H1 (ADRETUT06), 3111638H1 (BRSTNOT17)
17	44	2876494	THYRN0T10	619236R6 (PGANN0T01), 2876494F6 (THYRN0T10), 2876494H1 (THYRN0T10), 5875710H1 (BRAUN0T01), SCMA05016V1
18	45	3403225	ESOGNOT03	859412R1 (BRAITUT03), 859412T1 (BRAITUT03), 2618006H1 (GBLANOT01), 3403225F6 (ESOGNOT03), 3403225H1 (ESOGNOT03), 305556R6 (HEARN0T01), 1988284R6 (LUNGAST01), 2967845H1 (SCORN0T04), 3322082H1 (PTHYN0T03), 3495233H1 (ADRETUT07), 4163943X300V1 (BRSTNOT32), 4691781H1 (BRAENOT02), 5423053H1 (PROSTMT07)
19	46	4163943	BRSTNOT32	
20	47	4293484	BRABDIR01	570083R6 (NMLR3DT01), 3775144F6 (BRSTNOT27), 3841966H1 (DENDNOT01), 3845372H1 (DENDNOT01), 4293484H1 (BRABDIR01), 1283458T6 (COLNNOT16), 1671936F6 (BLADNOT05), 1671936T6 (BLADNOT05), 4174920H1 (SINTNOT21), 4901660H1 (OVARIDIT01), SAQA00885F1
21	48	4440080	SINTNOT22	
22	49	5495687	BRABDIR01	065572H1 (PLACNOB01), 364481R6 (PROSNOT01), 964976T1 (BRSTNOT05), 2204888F6 (SPLNFET02), 5495687H1 (BRABDIR01), 1217512T1 (NEUTGMT01), 1575649F1 (LNODNOT03), 2364831T6 (ADREN0T07), 3144524H1 (HNT2AZS07), SBZA04875V1, SBZA03722V1, SBZA05758V1
23	50	5527735	KIDNNOT34	
24	51	5540437	KIDNFEC01	5540437H1 (KIDNFEC01), 703559H1 (SYNORAT04), 898596R1 (BRSTTUT03), 1302747F6 (PLACNOT02), 1318515F1 (BLADNOT04), 1555032X12C1 (BLADTUT04), 1555177X14C1 (BLADTUT04), 1996026R6 (BRSTTUT03), 3356064H1 (PROSTTUT16)

Table 1 (cont.)

Polypeptide SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
25	52	5596281	COLCDIT03	985513R1 (LVENNOT03), 2862380F6 (SININOT03), 4255173F6 (BSCNNOT03), 4255173T6 (BSCNNOT03), 5596281H1 (COLCDIT03)
26	53	5731013	KIDCTMT01	1451660F6 (PENITUT01), 1495213T1 (PROSNON01), 2131940R6 (OVARNOT03), 2728885H1 (OVARUT05), 5731013H1 (KIDCTMT01)
27	54	5731162	KIDCTMT01	1229428X19 (BRAITUT01), 1238311X14R1 (LUNGUTUT02), 1257633F1 (MENITUT03), 1810038T6 (PROSTUT12), 3591284H1 (293TF5T01), 5731162H1 (KIDCTMT01), 5866723H1 (COLTDT04)

Table 2

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs and Domains	Homologous Sequences	Analytical Methods and Databases
1	468	S28 T34 T192 T273 S283 S299 T410 S23 S84 S160 S327 T414 Y141	N222	Signal peptide: M1-Pl6 ubiH/CO6 monooxygenase: D37-M50, K202-L236, A362-L389 Oxidoreductase motif: D37-S55, D37-S295, D339-A466 Aromatic ring hydroxylase motif: D37-H59, Q200-R215, R358- L389 Flavoprotein family: G255-F297, P332-L424 V335-L424	Predicted VISC ubiquinone monooxygenase [C. elegans] g2088820	BLAST-GenBank SPSCAN BLIMPS-BLOCKS BLIMPS-PRINTS BLAST-DOMO BLAST-PRODOM MOTIFS
2	254	T140 S163 T184 T4 T31 S71 T219		NifU family: S163-V241 NifU signature: V202-L220	Nitrogen fixation Nif U-like protein [A. thaliana] g4538920	BLAST-GenBank BLIMPS-PFAM BLAST-DOMO MOTIFS
3	555	S6 S10 T52 T103 T172 S213 S490 S66 T190 S270 T351 T365 S506 T521	N148	Signal peptide: M1-A40 D-amino acid oxidase: R26-A38 Aromatic ring hydroxylase motif: R26-F48 Flavin-containing amine oxidase signature: R26-E45 Flavin-containing monooxygenase motif: R26-K41	putative Cs protein with homology to polyamine oxidase [A. thaliana] g5123566	BLAST-GenBank SPSCAN BLIMPS-BLOCKS BLIMPS-PRINTS MOTIFS

Table 2 (cont.)

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs and Domains	Homologous Sequences	Analytical Methods and Databases
4	337	S337 T54 S142 S151 S171 S173 T12 T148 S301 S305 S322	N69	Signal peptide: M1-R36 Dihydroorotate dehydrogenase motif: L202-H211		SPSCAN BLIMPS-BLOCKS MOTIFS
5	109	S75		Signal peptide: M1-A55 Lipase/Acylhydrolase motif: D87-N92 beta-hydroxylase motif: G7-L101	beta-hydroxylase [Streptomyces verticillius] g507319	BLAST-GenBank SPSCAN BLIMPS-PFAM BLAST-PRODOM MOTIFS
6	385	S355 T37 S55 T99 T218 T275 S337 S346 S15 T54 S82 T190 S354		Oxidoreductase motif: V66-S110	putative oxidoreductase [Streptomyces coelicolor A3(2)] g3218376	BLAST-GenBank BLIMPS-PRODOM MOTIFS
7	312	T48 T266 S25 T32 S44 S77 S80 S91 S116 T209 S215	N218	Signal peptide: M1-G17 Oxidoreductase FAD/NAD binding domain: S169-G295, L279-P287, F90-W311 Phenolhydroxylase reductase family: F92-G104, V173-A192, L279-P287 Molybdopterin oxidoreductase family: S75-P236	phenolhydroxylase component [Acinetobacter calcoaceticus] g535285	BLAST-GenBank SPSCAN BLIMPS-PRINTS HMMER-PFAM BLAST-PRODOM BLAST-DOMO MOTIFS

Table 2 (cont.)

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs and Domains	Homologous Sequences	Analytical Methods and Databases
8	160	S22 S47 T65 T44 S47 Y157		Estradiol ring-cleavage dioxygenase proteins: D39-M61, T133-V154 Lipoxygenase Fe-binding region proteins: W23-D39 Mitochondrial P450 signature: C33-T44	biphenyl-2,3-diol 1,2-dioxygenase III-related protein [Vibrio g9657727	BLIMPS-BLOCKS BLIMPS-PRINTS MOTIFS
9	487	T237 S9 T24 T216 S248 T284 S342 T396 S404 S478 S9 S169 S196 S384	N167 N215 N394 N476	Cytochrome b5 heme binding domain family: E22-P99, H31-H78, V45-H55, H55-D69, T323-P344, A364-E407 Cytochrome b5 reductase signature: L269-K280, K290-G297, F360-L379, D398-L409, L455-P463 Oxidoreductase FAD binding domain: L356-L472 Eukaryotic molybdopterin reductase proteins: P56-R94, D381-E407 K451-G468, C244-H483, K249-P461	predicted oxidoreductase [C. elegans] g3881161	BLAST-GenBank HMMER-PFAM PROFILESAN BLIMPS-BLOCKS BLIMPS-PRINTS BLAST-PRODOM BLAST-DOMO MOTIFS

Table 2 (cont.)

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs and Domains	Homologous Sequences	Analytical Methods and Databases
10	524	T277 T68 S139 S187 T224 S305 S314 T106 S186 S388	N112 N168	Signal peptide: M1-A32 Transmembrane domain: M16-L35 Cytochrome P450: M1-Q49, A113-L512, L48-256, P52-L519, Y440-R490, F458-F489 E-class P450 group II signature: G141-K161, L197-Q215, D317-K363, Q377-F397, G417-E448, P455-F491 E-class P450 group IV signature: L378-P394, H428-D446, C468-L486 Mitochondrial P450 signature: G328-A345, R346-Q359, A376-P394, I459-K479	leukotriene B4 omega hydroxylase [Homo sapiens] g1857022	BLAST-GenBank HMMER SPSCAN HMMER-PFAM PROFILES CAN BLIMPS-BLOCKS BLIMPS-PRINTS BLAST-PRODOM BLAST-DOMO MOTIFS
11	144	S12 S58 T127		Signal peptide: M1-C13 Glutaredoxin proteins: V2-L20 Anion exchanger: A39-I110 Fungal Zn/Cys binuclear cluster signature: S9-K15	predicted arsenate reductase [Bacillus subtilis] g2635777	BLAST-GenBank SPSCAN HMMER-PFAM PROFILES CAN BLIMPS-PRINTS MOTIFS

Table 2 (cont.)

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs and Domains	Homologous Sequences	Analytical Methods and Databases
12	373	T238 S268 T330 S355 T137 S225 S360 S366	N118 N190	Signal peptide: M1-S32 Zn alcohol dehydro- genase signature: A36-T372, D72-N88, K56-L340, V73-A359, S85-T137, P101-V128 Shared mitochondrial respiratory function and transcription factor motif: R44-P343, S35-D352	nuclear receptor binding factor with homology to Zn-binding dehydrogenase NRBF-1 [Rattus norvegicus] g3970880 Masuda, N. et al. (1998) Gene 221:225-233	BLAST-GenBank SPSCAN PROFILES-SCAN HMMER-PFAM BLAST-DOMO BLAST-PRODOM BLIMPS-BLOCKS MOTIFS
13	305	T152 T166 T199 T218 T224 T38 S58 S250		malate dehydrogenase: P131-T300 NADH-Ubiquinone/ plastoquinone: K126-T137 Phthalate dioxygenase reductase signature: F223-R232	malate dehydrogenase [Echinococcus granulosus] g3386331	BLAST-GenBank BLAST-DOMO BLIMPS-PFAM BLIMPS-PRINTS MOTIFS
14	500	S279 T24 S136 T183 S226 T259 S349 S394 S432 T465 S483 T194 S252 Y127	N480	Signal peptide: M1-C25 Leucine zipper: L467-L488 GMC oxidoreductases: K12-A30 FAD-dependent pyridine nucleotide and class-I disulphide nucleotide reductase signature: K12-P34, D311-P325, I353-C360	predicted oxidoreductase [C. elegans] g3874510	BLAST-GenBank SPSCAN BLIMPS-BLOCKS BLIMPS-PRINTS MOTIFS

Table 2 (cont.)

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs and Domains	Homologous Sequences	Analytical Methods and Databases
15	369	S138 S52 T117 S335 S24 T250 S256 Y155	N211	Beta-hydroxylase: V188-I364	beta-hydroxylase [Streptomyces verticillius] g507319	BLAST-GenBank BLAST-PRODOM MOTIFS
16	145	T73 S92 T99 T104 T129 T130	N116		NADH:ubiquinone oxidoreductase b17.2 subunit [Bos taurus] g4006932	BLAST-GenBank MOTIFS
17	255	T61 S129 T130 S193 T25 T76 T120 T157	N45	Signal peptide: M1-T25; M1-A30 Transmembrane domain: G229-L249 PEP-utilizing enzyme signature: L228-G235		HMNER SPSCAN BLIMPS-BLOCKS MOTIFS
18	246	S21 T55 S104 T187 S208 S221	N227 N235	Transmembrane domain: M1-A19 Short chain ADH family: V43-L241, T33-P222, K117-V128 Glucose/ribitol dehydrogenase family: V41-E58, K117-V128	androgen- regulated short- chain dehydrogenase/red uctase 1 [Homo sapiens] g9622124	BLAST-GenBank HMNER HMNER-PFAM BLIMPS-PRINTS BLAST-DOMO MOTIFS
19	467	T105 S118 S141 T190 S239 S426 S452 S39 T80 T145 T212 T393 T406 S463	N83	Fe-ADH family: V51-T253, S52-I336, G57-Y251, S264-M464, V184-E193, I307-E378, G274-L456 ATP/GTP binding site (P loop): G35-T42	putative type III alcohol dehydrogenase [D. melanogaster] g2431772	BLAST-GenBank HMNER-PFAM BLIMPS-BLOCKS PROFILES-SCAN BLAST-PRODOM MOTIFS

Table 2 (cont.)

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs and Domains	Homologous Sequences	Analytical Methods and Databases
20	317	S76 T97 S215 S297 T283 Y99	N276	Leucine zipper: L246-L267 3-beta-hydroxysteroid dehydrogenase: M1-T313, V11-E33, G56-Q310, H88-T140, A170-R214, Y270-Q317 UDP-glucose epimerase: L10-D43, G56-Q310	type IV beta- hydroxysteroid dehydrogenase CCA2 [Rattus norvegicus] g2563999 Hayashi, Y. et al. (1997) Biochim. Biophys. Acta 1352:145-150	BLAST-GenBank HMMER-PFAM BLIMPS-PFAM BLAST-DOMO BLAST-PRODOM MOTIFS
21		S10 S56		Signal peptide: M1-A53 Glucose/ribitol dehydrogenase family: S10-L21, E59-G75, Y87-R106, T110-A127, E143-T163 Short chain ADH family: M1-Q125, S10-G20, G65-M73, S74-R102, Y87-R106, I67-E104, H111-G120 G65-G120, I5-T163	antenna-specific short-chain dehydrogenase/ reductase [D. melanogaster] g4530425	BLAST-GenBank SPSCAN HMMER-PFAM BLIMPS-PRINTS BLIMPS-BLOCKS PROFILES-SCAN BLAST-DOMO MOTIFS
22	361	T86 T149 T183 T295 T306 S39 S169 S172 T189 S254 T280 T286 T341 S346	N165 N181 N187 N194 N206 N278 N293	Signal peptide: M1-G32; M1-V35 Transmembrane domain: D323-I342 Thioredoxin family: T211-W219, W219-P228 E-class P450 group II signature: H66-G83	thioredoxin [Fasciola hepatica] g6492215	HMMER SPSCAN BLIMPS-PRINTS MOTIFS

Table 2 (cont.)

Polypeptide SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs and Domains	Homologous Sequences	Analytical Methods and Databases
23	477	T140 T12 S67 T79 S90 S161 S165 S254 T273 S287 S304 T449 T58 T135 T211 S310 T340 Y49		Fe hydrogenase family: Q89-Q405, A104-Q405 Four disulfide bridge: K381-P386	contains similarity to hydrogenases [Caenorhabditis elegans] g7332095	BLAST-GenBank BLAST-PRODOM BLAST-DOMO BLIMPS-BLOCKS MOTIFS
24	621	S337 S602 S60 S185 T197 S237 T462 S493 T522 S584 S84 S191 S302 T328 T449 S553 Y506	N208	Acyl-CoA dehydrogenase signature: R85-L438, E130-A437, L104-E114, Y205-G217, G253-F293, M307-E357, E396-L438, L180-G236, A379-I431, R37-L155, E52-G440, Q399-D418	very-long-chain acyl-CoA dehydrogenase [Mus musculus] g2765125	BLAST-GenBank HMMER-PFAM PROFILES-SCAN BLIMPS-BLOCKS BLAST-PRODOM BLAST-DOMO MOTIFS
25	246	S70 T89 S184 T213 T60 S137 T149 T208 Y49	N39 N130	Signal peptide: M1-A21 Glucose/ribitol dehydrogenase family: V8-A25, E73-F84, M120-S136, Y147-I166, Q168-L185, K207-D227 Oxidoreductase domain: V146-G242 Short chain ADH family: M1-P183, L4-W243, E73-F84, G126-G178, Y147-I166	3-oxoacyl-(acyl carrier protein) reductase [Thermotoga maritima] g4982301	BLAST-GenBank SPSCAN HMMER-PFAM BLIMPS-BLOCKS PROFILES-SCAN BLIMPS-PRINTS BLAST-PRODOM BLAST-DOMO MOTIFS
26	160	S93 S108 T2 S39 T134 T139		Signal peptide: M1-G20 Thioesterase domain proteins: R114-I126	predicted NADH- ubiquinone oxidoreductase B8 subunit [C. elegans] g3874440	BLAST-GenBank SPSCAN BLIMPS-PFAM MOTIFS
27	292	S25 S54 T55 S72 S120 S199 S253	N251	Modulation hydrolase: V232-G242		BLIMPS-PRODOM MOTIFS

Table 3

Nucleotide SEQ ID NO:	Selected Fragments	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
28		Reproductive (0.341) Nervous (0.136) Gastrointestinal (0.114) Urologic (0.114)	Cancer (0.477) Cell proliferation (0.182) Inflammation (0.136) Trauma (0.136)	PSPORT1
29	1-215	Reproductive (0.237) Gastrointestinal (0.184) Cardiovascular (0.171)	Cancer (0.592) Cell proliferation (0.171) Trauma (0.105)	PSPORT1
30	1-259 416-1039	Nervous (0.294) Reproductive (0.176) Cardiovascular (0.137)	Cancer (0.373) Inflammation (0.235) Cell proliferation (0.176)	pINCY
31	229-315	Hematopoietic/Immune (0.250) Nervous (0.250) Gastrointestinal (0.125) Reproductive (0.125)	Inflammation (0.562) Cancer (0.438) Cell proliferation (0.188)	pINCY
32	427-498	Nervous (0.667) Hematopoietic/Immune (0.111)	Cancer (0.333) Inflammation (0.333) Cell proliferation (0.278)	PSPORT1
33	314-357 428-481 911-1024	Nervous (0.256) Gastrointestinal (0.154) Reproductive (0.128)	Cancer (0.436) Inflammation (0.256) Cell proliferation (0.231)	pINCY
34	194-664 1011-1134	Reproductive (0.231) Nervous (0.231) Hematopoietic/Immune (0.154) Endocrine (0.154)	Cancer (0.462) Trauma (0.231) Cell proliferation (0.154)	pINCY
35	1-177	Gastrointestinal (0.750) Reproductive (0.250)	Cancer (0.250) Cell proliferation (0.250) Trauma (0.250)	pINCY
36	645-911	Gastrointestinal (0.250) Reproductive (0.250) Hematopoietic/Immune (0.167)	Cancer (0.528) Inflammation (0.306) Trauma (0.111)	pINCY

Table 3 (cont.)

Nucleotide SEQ ID NO:	Selected Fragments	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
37		Gastrointestinal (0.405) Reproductive (0.262) Urologic (0.119)	Cancer (0.524) Inflammation (0.262)	pINCY
38	1-79 173-215 440-549	Nervous (1.000)	Trauma (1.000)	PSPORT1
39		Reproductive (0.300) Gastrointestinal (0.160) Nervous (0.120)	Cancer (0.420) Inflammation (0.280) Cell proliferation (0.200)	pINCY
40	1-304 572-1196	Cardiovascular (0.143) Developmental (0.143) Endocrine (0.143) Gastrointestinal (0.143) Reproductive (0.143) Urologic (0.143) Nervous (0.143)	Trauma (0.429) Cell proliferation (0.286) Inflammation (0.143) Neurological (0.143)	pINCY
41	448-765 1039-1926	Reproductive (0.308) Nervous (0.154) Cardiovascular (0.115) Hematopoietic/Immune (0.115) Gastrointestinal (0.115)	Cancer (0.538) Cell proliferation (0.192) Inflammation (0.192)	pINCY
42	490-849 1072-1152	Nervous (0.538) Hematopoietic/Immune (0.154) Reproductive (0.154)	Cancer (0.385) Cell proliferation (0.154) Inflammation (0.154) Trauma (0.154)	PSPORT1
43	1-38	Reproductive (0.316) Gastrointestinal (0.137) Cardiovascular (0.137)	Cancer (0.463) Inflammation (0.232) Cell proliferation (0.168)	pINCY
44	243-539 555-1352	Nervous (0.500) Endocrine (0.167) Musculoskeletal (0.167) Gastrointestinal (0.167)	Cancer (0.333) Inflammation (0.333) Neurological (0.167) Trauma (0.167)	pINCY

Table 3 (cont.)

Nucleotide SEQ ID NO:	Selected Fragments	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
45	127-171	Gastrointestinal (0.500) Reproductive (0.250) Nervous (0.250)	Cancer (0.750) Inflammation (0.250)	pINCY
46	1-80 300-338 792-833	Reproductive (0.229) Nervous (0.171) Cardiovascular (0.143)	Cancer (0.400) Inflammation (0.257) Trauma (0.171)	pINCY
47	465-527 612-692 810-968	Nervous (0.211) Reproductive (0.211) Gastrointestinal (0.158) Hematopoietic/Immune (0.158)	Cancer (0.474) Inflammation (0.368) Cell proliferation (0.105)	pINCY
48	655-1305	Gastrointestinal (0.412) Reproductive (0.206) Nervous (0.118)	Cancer (0.471) Inflammation (0.235) Trauma (0.206)	PBLUESC RIPT
49	1-69 1397-1459	Reproductive (0.361) Nervous (0.194) Developmental (0.111)	Cancer (0.389) Inflammation (0.250) Cell proliferation (0.222)	pINCY
50	1-51 157-198 1246-2101	Reproductive (0.235) Nervous (0.216) Hematopoietic/Immune (0.157)	Cancer (0.353) Inflammation (0.275) Cell proliferation (0.255)	pINCY
51	1005-1073 1233-1328	Reproductive (0.247) Nervous (0.195) Gastrointestinal (0.143)	Cancer (0.442) Inflammation (0.273) Cell proliferation (0.156)	pINCY
52	1-82	Reproductive (0.255) Nervous (0.181) Gastrointestinal (0.138)	Cancer (0.447) Inflammation (0.191) Cell proliferation (0.138)	pINCY
53	1-94	Reproductive (0.327) Nervous (0.143) Gastrointestinal (0.102)	Cancer (0.510) Cell proliferation (0.204) Inflammation (0.204)	pINCY
54	1-469 998-1624	Reproductive (0.328) Nervous (0.197) Cardiovascular (0.180)	Cancer (0.639) Inflammation (0.213)	pINCY

Table 4

Nucleotide SEQ ID NO:	Library	Library Description
28	OVARNOT02	This library was constructed using RNA isolated from ovarian tissue from a 59-year-old Caucasian female who died of a myocardial infarction. Patient history included cardiomyopathy, coronary artery disease, previous myocardial infarctions, hypercholesterolemia, hypotension, and arthritis.
29	COLNNOT09	This library was constructed using RNA isolated from colon tissue from a 60-year-old Caucasian male.
30	BRAINOT11	This library was constructed using RNA isolated from brain tissue from the right temporal lobe of a 5-year-old Caucasian male. Pathology indicated extensive polymicrogyria and mild to moderate gliosis (predominantly subpial and subcortical), consistent with chronic seizure disorder. Family history included a cervical neoplasm.
31	PLACNOT02	This library was constructed using RNA isolated from the placental tissue of a Hispanic female fetus, who was prematurely delivered at 21 weeks' gestation. Serology of the mother's blood was positive for CMV (cytomegalovirus).
32	SINTTUT01	This library was constructed using RNA isolated from small intestine tumor tissue from a 42-year-old Caucasian male. Carcinoid tumor was identified in the ileum. Patient history included benign hypertension. Family history included benign hypertension, a cerebrovascular accident, malignant neoplasm of prostate, and tuberculosis.
33	BRAINOT14	This library was constructed using RNA isolated from brain tissue from the left frontal lobe of a 40-year-old Caucasian female. Pathology for the associated tumor tissue indicated gemistocytic astrocytoma.
34	PROSTUT10	This library was constructed using RNA isolated from prostatic tumor tissue from a 66-year-old Caucasian male. Pathology indicated an adenocarcinoma. Adenofibromatous hyperplasia was also present. The patient presented with elevated prostate specific antigen (PSA). Family history included prostate cancer and secondary bone cancer.
35	PROSNOT19	This library was constructed using RNA isolated from diseased prostate tissue from a 59-year-old Caucasian male. Pathology indicated adenofibromatous hyperplasia. Pathology for the associated tumor tissue indicated an adenocarcinoma. The patient presented with elevated prostate-specific antigen (PSA). Patient history included colon diverticuli, asbestosis, and thrombophlebitis. Family history included benign hypertension, multiple myeloma, hyperlipidemia and rheumatoid arthritis.

Table 4 (cont.)

Nucleotide SEQ ID NO:	Library	Library Description
36	PANCTUT02	This library was constructed using RNA isolated from pancreatic tumor tissue from a 45-year-old Caucasian female. Pathology indicated anaplastic carcinoma. Family history included benign hypertension, hyperlipidemia and atherosclerotic coronary artery disease.
37	LIVRFET02	This library was constructed using RNA isolated from liver tissue from a Caucasian female fetus, who died at 20 weeks' gestation. Family history included seven days of erythromycin treatment for bronchitis in the mother during the first trimester.
38	BSTMNON02	This normalized brain stem library was constructed from 2.84 million independent clones from a brain stem library. Starting RNA was made from the brain stem tissue of a 72-year-old Caucasian male who died from myocardial infarction. Patient history included coronary artery disease, insulin-dependent diabetes mellitus, and arthritis. Normalization and hybridization conditions were adapted from Soares et al. (PNAS (1994) 91:9228).
39	CONUTUT01	This library was constructed using RNA isolated from sigmoid mesentery tumor tissue from a 61-year-old female. Pathology indicated a metastatic malignant mixed mullerian tumor present in the sigmoid mesentery at two sites.
40	BRAINOT18	This library was constructed using RNA isolated from left temporal lobe brain tissue from a 34-year-old Caucasian male. Pathology for the associated tumor tissue indicated metastatic malignant melanoma. Neoplastic cells strongly expressed HMB-45. Patient history included malignant melanoma of skin of the trunk. Family history included liver cancer, acute myocardial infarction, atherosclerotic coronary artery disease, and cerebrovascular disease.
41	OVARTUT03	This library was constructed using RNA isolated from ovarian tumor tissue from the left ovary of a 52-year-old mixed ethnicity female. Pathology indicated an invasive seroanaplastic carcinoma forming a mass in the left ovary. Multiple tumor implants were present on the surface of both ovaries and fallopian tubes and the uterus. Pathology also indicated a metastatic seroanaplastic carcinoma involving the omentum, cul-de-sac peritoneum, left broad ligament peritoneum, and mesentery colon. Patient history included breast cancer, chronic peptic ulcer, and joint pain. Family history included colon cancer, cerebrovascular disease, breast cancer, type II diabetes, esophagus cancer, and depressive disorder.

Table 4 (cont.)

Nucleotide SEQ ID NO:	Library	Library Description
42	THP1AZS08	This subtracted library was constructed from a 5-aza-deoxycytidine treated THP-1 cell library. Starting RNA was made from THP-1 promonocyte cells treated for three days with 0.8 micromolar AZ. The hybridization probe for subtraction was derived from a similarly constructed library, made from RNA isolated from untreated THP-1 cells. 5.76 million clones from the AZ-treated THP-1 cell library were then subjected to two rounds of subtractive hybridization with 5 million clones from the untreated THP-1 cell library. Subtractive hybridization conditions were based on the methodologies of Swaroop et al. NAR (1991) 19:1954, and Bonaldo et al. (1996) Genome Research 6:791. THP-1 (ATCC TIB 202) is a human promonocyte line derived from peripheral blood of a 1-year-old Caucasian male with acute monocytic leukemia.
43	ADRETUT06	This library was constructed using RNA isolated from adrenal tumor tissue from a 57-year-old Caucasian female. Pathology indicated pheochromocytoma, forming a nodular mass completely replacing the medulla of the adrenal gland.
44	THYRNOT10	This library was constructed using RNA isolated from the diseased left thyroid tissue from a 30-year-old Caucasian female. Pathology indicated lymphocytic thyroiditis.
45	ESOGNOT03	This library was constructed using RNA isolated from esophageal tissue from a 53-year-old Caucasian male. Patient history included membranous nephritis, hyperlipidemia, benign hypertension, and anxiety state. Family history included atherosclerotic coronary artery disease, cirrhosis, an abdominal aortic aneurysm rupture, breast cancer, and myocardial infarction.
46	BRSTNOT32	This library was constructed using RNA isolated from diseased breast tissue from a 46-year-old Caucasian female. Pathology indicated nonproliferative fibrocystic disease. Family history included breast cancer, benign hypertension, and atherosclerotic coronary artery disease.
47	BRABDIR01	This library was constructed using RNA isolated from diseased cerebellum tissue from the brain of a 57-year-old Caucasian male, who died from a cerebrovascular accident.
48	SINTNOT22	This library was constructed using RNA isolated from small intestine tissue from a 15-year-old Caucasian female who died from a closed head injury. Serology was positive for cytomegalovirus. Patient history included seasonal allergies and marijuana use.

Table 4 (cont.)

Nucleotide SEQ ID NO:	Library	Library Description
49	BRABDIR01	This library was constructed using RNA isolated from diseased cerebellum tissue from the brain of a 57-year-old Caucasian male, who died from a cerebrovascular accident.
50	KIDNNOT34	This library was constructed using RNA isolated from left kidney tissue obtained from an 8-year-old Caucasian male who died from an intracranial hemorrhage. Serologies were negative.
51	KIDNFEC01	This library was constructed using RNA isolated from kidney tissue from a pool of twelve Caucasian male and female fetuses that were spontaneously aborted at 19-23 weeks' gestation.
52	COLCDIT03	This library was constructed using RNA isolated from diseased colon polyp tissue from the cecum of a 67-year-old female. Pathology indicated a benign cecum polyp. Pathology for the associated tumor tissue indicated invasive grade 3 adenocarcinoma that arose in tubulovillous adenoma forming a fungating mass in the cecum.
53	KIDCTMT01	This library was constructed using RNA isolated from kidney cortex tissue from a 65-year-old male. Pathology for the associated tumor tissue indicated renal cell carcinoma within the mid-portion of the kidney and the renal capsule.
54	KIDCTMT01	This library was constructed using RNA isolated from kidney cortex tissue from a 65-year-old male. Pathology for the associated tumor tissue indicated renal cell carcinoma within the mid-portion of the kidney and the renal capsule.

Table 5

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25:3389-3402.	ESTs: Probability value= 1.0E-8 or less Full Length sequences: Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises at least five functions: fasta, tfasta, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad. Sci. USA 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183:63-98; and Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489.	ESTs: fasta E value=1.0E-6 Assembled ESTs: fasta Identity= 95% or greater and Match length=200 bases or greater; fastx E value=1.0E-8 or less Full Length sequences: fastx score=100 or greater
BLIMPS	A BLOCKS IMProved Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S. and J.G. Henikoff (1991) Nucleic Acids Res. 19:6565-6572; Henikoff, J.G. and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37:417-424.	Probability value= 1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) J. Mol. Biol. 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322; Durbin, R. et al. (1998) Our World View, in a Nutshell, Cambridge Univ. Press, pp. 1-350.	Probability value= 1.0E-3 or less for PFAM hits

Table 5 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, M. et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221.	Normalized quality score \geq GCG-specified "HIGH" value for that particular Prosite motif. Generally, score=1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M.S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score= 120 or greater; Match length= 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies.	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12:431-439.	Score=3.5 or greater
TMAP	A program that uses weight matrices to delineate transmembrane segments on protein sequences and determine orientation.	Persson, B. and P. Argos (1994) J. Mol. Biol. 237:182-192; Persson, B. and P. Argos (1996) Protein Sci. 5:363-371.	
TMHMMER	A program that uses a hidden Markov model (HMM) to delineate transmembrane segments on protein sequences and determine orientation.	Sonnhammer, E.L. et al. (1998) Proc. Sixth Intl. Conf. on Intelligent Systems for Mol. Biol., Glasgow et al., eds., The Am. Assoc. for Artificial Intelligence Press, Menlo Park, CA, pp. 175-182.	
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

What is claimed is:

1. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of:

- 5 a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-27,
 b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-27,
 c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-27, and
10 d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-27.

2. An isolated polypeptide of claim 1 selected from the group consisting of SEQ ID NO:1-27.

15 3. An isolated polynucleotide encoding a polypeptide of claim 1.

4. An isolated polynucleotide encoding a polypeptide of claim 2.

20 5. An isolated polynucleotide of claim 4 selected from the group consisting of SEQ ID NO:28-54.

6. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 3.

25 7. A cell transformed with a recombinant polynucleotide of claim 6.

8. A transgenic organism comprising a recombinant polynucleotide of claim 6.

30 9. A method for producing a polypeptide of claim 1, the method comprising:

 a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide, and said recombinant polynucleotide comprises a promoter sequence operably linked to a polynucleotide encoding the polypeptide of claim 1, and

35 b) recovering the polypeptide so expressed.

10. An isolated antibody which specifically binds to a polypeptide of claim 1.

11. An isolated polynucleotide comprising a polynucleotide sequence selected from the group consisting of:

- 5 a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:28-54,
 b) a naturally occurring polynucleotide sequence having at least 70% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:28-54,
 c) a polynucleotide sequence complementary to a),
 d) a polynucleotide sequence complementary to b), and
10 e) an RNA equivalent of a)-d).

12. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 11.

15 13. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, the method comprising:

- a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization
20 complex is formed between said probe and said target polynucleotide or fragments thereof, and
 b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.

14. A method of claim 13, wherein the probe comprises at least 60 contiguous nucleotides.

25 15. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, the method comprising:

- a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and
30 b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

16. A composition comprising an effective amount of a polypeptide of claim 1 and a pharmaceutically acceptable excipient.

17. A composition of claim 16, wherein the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-27.

18. A method for treating a disease or condition associated with decreased expression of functional ORP, comprising administering to a patient in need of such treatment the composition of claim 16.

19. A method for screening a compound for effectiveness as an agonist of a polypeptide of claim 1, the method comprising:

- 10 a) exposing a sample comprising a polypeptide of claim 1 to a compound, and
- b) detecting agonist activity in the sample.

20. A composition comprising an agonist compound identified by a method of claim 19 and a pharmaceutically acceptable excipient.

15 21. A method for treating a disease or condition associated with decreased expression of functional ORP, comprising administering to a patient in need of such treatment a composition of claim 20.

20 22. A method for screening a compound for effectiveness as an antagonist of a polypeptide of claim 1, the method comprising:

- a) exposing a sample comprising a polypeptide of claim 1 to a compound, and
- b) detecting antagonist activity in the sample.

25 23. A composition comprising an antagonist compound identified by a method of claim 22 and a pharmaceutically acceptable excipient.

24. A method for treating a disease or condition associated with overexpression of functional ORP, comprising administering to a patient in need of such treatment a composition of claim 23.

30 25. A method of screening for a compound that specifically binds to the polypeptide of claim 1, said method comprising the steps of:

- a) combining the polypeptide of claim 1 with at least one test compound under suitable conditions, and
- 35 b) detecting binding of the polypeptide of claim 1 to the test compound, thereby identifying a

compound that specifically binds to the polypeptide of claim 1.

26. A method of screening for a compound that modulates the activity of the polypeptide of claim 1, said method comprising:

- 5 a) combining the polypeptide of claim 1 with at least one test compound under conditions permissive for the activity of the polypeptide of claim 1,
- b) assessing the activity of the polypeptide of claim 1 in the presence of the test compound, and
- c) comparing the activity of the polypeptide of claim 1 in the presence of the test compound
- 10 with the activity of the polypeptide of claim 1 in the absence of the test compound, wherein a change in the activity of the polypeptide of claim 1 in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide of claim 1.

27. A method for screening a compound for effectiveness in altering expression of a target

15 polynucleotide, wherein said target polynucleotide comprises a sequence of claim 5, the method comprising:

- a) exposing a sample comprising the target polynucleotide to a compound, under conditions suitable for the expression of the target polynucleotide,
- b) detecting altered expression of the target polynucleotide, and
- 20 c) comparing the expression of the target polynucleotide in the presence of varying amounts of the compound and in the absence of the compound.

28. A method for assessing toxicity of a test compound, said method comprising:

- a) treating a biological sample containing nucleic acids with the test compound;
- 25 b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide of claim 11 under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide comprising a polynucleotide sequence of a polynucleotide of claim 11 or fragment thereof;
- 30 c) quantifying the amount of hybridization complex; and
- d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

SEQUENCE LISTING

<110> INCYTE GENOMICS, INC.

YUE, Henry
 LAL, Preeti
 TANG, Y. Tom
 HILLMAN, Jennifer
 BAUGHN, Mariah R.
 AZIMZAI, Yalda
 LU, Dyung Aina M.

<120> HUMAN OXIDOREDUCTASE PROTEINS

<130> PF-0754 PCT

<140> To Be Assigned

<141> Herewith

<150> 60/172,367

<151> 1999-12-16

<160> 54

<170> PERL Program

<210> 1

<211> 468

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 543496CD1

<400> 1

Met	Ala	Ala	Arg	Leu	Val	Ser	Arg	Cys	Gly	Ala	Val	Arg	Ala	Ala		
1				5					10					15		
Pro	His	Ser	Gly	Pro	Leu	Val	Ser	Trp	Arg	Arg	Trp	Ser	Gly	Ala		
				20					25					30		
Ser	Thr	Asp	Thr	Val	Tyr	Asp	Val	Val	Val	Ser	Gly	Gly	Gly	Leu		
				35					40					45		
Val	Gly	Ala	Ala	Met	Ala	Cys	Ala	Leu	Gly	Tyr	Asp	Ile	His	Phe		
				50					55					60		
His	Asp	Lys	Lys	Ile	Leu	Leu	Leu	Glu	Ala	Gly	Pro	Lys	Lys	Val		
				65					70					75		
Leu	Glu	Lys	Leu	Ser	Glu	Thr	Tyr	Ser	Asn	Arg	Val	Ser	Ser	Ile		
				80					85					90		
Ser	Pro	Gly	Ser	Ala	Thr	Leu	Leu	Ser	Ser	Phe	Gly	Ala	Trp	Asp		
				95					100					105		
His	Ile	Cys	Asn	Met	Arg	Tyr	Arg	Ala	Phe	Arg	Arg	Met	Gln	Val		
				110					115					120		
Trp	Asp	Ala	Cys	Ser	Glu	Ala	Leu	Ile	Met	Phe	Asp	Lys	Asp	Asn		
				125					130					135		
Leu	Asp	Asp	Met	Gly	Tyr	Ile	Val	Glu	Asn	Asp	Val	Ile	Met	His		
				140					145					150		
Ala	Leu	Thr	Lys	Gln	Leu	Glu	Ala	Val	Ser	Asp	Arg	Val	Thr	Val		
				155					160					165		
Leu	Tyr	Arg	Ser	Lys	Ala	Ile	Arg	Tyr	Thr	Trp	Pro	Cys	Pro	Phe		
				170					175					180		
Pro	Met	Ala	Asp	Ser	Ser	Pro	Trp	Val	His	Ile	Thr	Leu	Gly	Asp		
				185					190					195		
Gly	Ser	Thr	Phe	Gln	Thr	Lys	Leu	Leu	Ile	Gly	Ala	Asp	Gly	His		
				200					205					210		
Asn	Ser	Gly	Val	Arg	Gln	Ala	Val	Gly	Ile	Gln	Asn	Val	Ser	Trp		

Asn Tyr Asp Gln	215	220	225
Ser Ala Val Val Ala	230	Thr Leu His Leu Ser	Glu
Ala Thr Glu Asn	245	235	240
Asn Val Ala Trp Gln	250	Arg Phe Leu Pro Ser	Gly
Pro Ile Ala Leu	260	Thr Leu Ser Ser Leu	Val
Trp Ser Thr Ser	275	265	270
His Glu His Ala Ala	280	Glu Leu Val Ser Met	Asp
Glu Glu Lys Phe	290	295	300
Val Asp Ala Val Asn	305	Ser Ala Phe Trp Ser	Asp
Ala Asp His Thr	320	Ala Gly Ala Met Leu	Gln
Tyr Ala Val Ser	335	310	315
Leu Leu Lys Pro Thr	350	Lys Val Ser Ala Arg	Gln
Leu Pro Pro Ser	365	325	330
Val Ala Arg Val Asp	380	Ala Lys Ser Arg Val	Leu
Phe Pro Leu Gly	395	340	345
Leu Gly His Ala Ala	410	Glu Tyr Val Arg Pro	Arg
Val Ala Leu Ile	425	355	360
Gly Asp Ala Ala His	440	Arg Val His Pro Leu	Ala
Gly Gln Gly Val	455	370	375
Asn Met Gly Phe Gly	460	385	390
His His Leu Ser	475	Asp Ile Ser Ser Leu	Ala
Thr Ala Ala Phe Asn	490	400	405
Gly Lys Asp Leu Gly	505	415	420
Val Ser His Leu Thr	520	Glu Arg Gln Arg His	Asn
Thr Ala Leu Leu	535	430	435
Ala Ala Thr Asp Leu	550	Leu Lys Arg Leu Tyr	Ser
Thr Ser Ala Ser	565	445	450
Pro Leu Val Leu Leu	580	Arg Thr Trp Gly Leu	Gln
Ala Thr Asn Ala	595	460	465
Val Ser Pro Leu Lys		Glu Gln Ile Met Ala	Phe
Ala Ser Lys			

<210> 2
 <211> 254
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 907607CD1

<400> 2
 Met Ala Ala Thr Ala Arg Arg Gly Trp Gly Ala Ala Ala Val Ala
 1 5 10 15
 Ala Gly Leu Arg Arg Arg Phe Cys His Met Leu Lys Asn Pro Tyr
 20 25 30
 Thr Ile Lys Lys Gln Pro Leu His Gln Phe Val Gln Arg Pro Leu
 35 40 45
 Phe Pro Leu Pro Ala Ala Phe Tyr His Pro Val Arg Tyr Met Phe
 50 55 60
 Ile Gln Thr Gln Asp Thr Pro Asn Pro Asn Ser Leu Lys Phe Ile
 65 70 75
 Pro Gly Lys Pro Val Leu Glu Thr Arg Thr Met Asp Phe Pro Thr
 80 85 90
 Pro Ala Ala Ala Phe Arg Ser Pro Leu Ala Arg Gln Leu Phe Arg
 95 100 105
 Ile Glu Gly Val Lys Ser Val Phe Phe Gly Pro Asp Phe Ile Thr
 110 115 120
 Val Thr Lys Glu Asn Glu Glu Leu Asp Trp Asn Leu Leu Lys Pro
 125 130 135
 Asp Ile Tyr Ala Thr Ile Met Asp Phe Phe Ala Ser Gly Leu Pro

Leu Val Thr Glu	140	Glu Thr Pro Ser Gly	145	Glu Ala Gly Ser Glu	150
Asp Asp Glu Val	155	Val Ala Met Ile Lys	160	Glu Leu Leu Asp Thr	165
Ile Arg Pro Thr	170	Val Gln Glu Asp Gly	175	Gly Asp Val Ile Tyr	180
Gly Phe Glu Asp	185	Gly Ile Val Gln Leu	190	Lys Leu Gln Gly Ser	195
Thr Ser Cys Pro	200	Ser Ser Ile Ile Thr	205	Leu Lys Asn Gly Ile	210
Asn Met Leu Gln	215	Phe Tyr Ile Pro Glu	220	Val Glu Gly Val Glu	225
Val Met Asp Asp	230	Glu Ser Asp Glu Lys	235	Glu Ala Asn Ser Pro	240
	245		250		

<210> 3

<211> 555

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1290078CD1

<400> 3

Met Gln Ser Cys Glu	5	Ser Ser Gly Asp Ser	10	Ala Asp Asp Pro Leu	15
Ser Arg Gly Leu Arg	20	Arg Arg Gly Gln Pro	25	Arg Val Val Val Ile	30
Gly Ala Gly Leu Ala	35	Gly Leu Ala Ala Ala	40	Lys Ala Leu Leu Glu	45
Gln Gly Phe Thr Asp	50	Val Thr Val Leu Glu	55	Ala Ser Ser His Ile	60
Gly Gly Arg Val Gln	65	Ser Val Lys Leu Gly	70	His Ala Thr Phe Glu	75
Leu Gly Ala Thr Trp	80	Ile His Gly Ser His	85	Gly Asn Pro Ile Tyr	90
His Leu Ala Glu Ala	95	Asn Gly Leu Leu Glu	100	Glu Thr Thr Asp Gly	105
Glu Arg Ser Val Gly	110	Arg Ile Ser Leu Tyr	115	Ser Lys Asn Gly Val	120
Ala Cys Tyr Leu Thr	125	Asn His Gly Arg Arg	130	Ile Pro Lys Asp Val	135
Val Glu Glu Phe Ser	140	Asp Leu Tyr Asn Glu	145	Val Tyr Asn Leu Thr	150
Gln Glu Phe Phe Arg	155	His Asp Lys Pro Val	160	Asn Ala Glu Ser Gln	165
Asn Ser Val Gly Val	170	Phe Thr Arg Glu Glu	175	Val Arg Asn Arg Ile	180
Arg Asn Asp Pro Asp	185	Asp Pro Glu Ala Thr	190	Lys Arg Leu Lys Leu	195
Ala Met Ile Gln Gln	200	Tyr Leu Lys Val Glu	205	Ser Cys Glu Ser Ser	210
Ser His Ser Met Asp	215	Glu Val Ser Leu Ser	220	Ala Phe Gly Glu Trp	225
Thr Glu Ile Pro Gly	230	Ala His His Ile Ile	235	Pro Ser Gly Phe Met	240
Arg Val Val Glu Leu	245	Leu Ala Glu Gly Ile	250	Pro Ala His Val Ile	255
Gln Leu Gly Lys Pro	260	Val Arg Cys Ile His	265	Trp Asp Gln Ala Ser	270
Ala Arg Pro Arg Gly	275	Pro Glu Ile Glu Pro	280	Arg Gly Glu Gly Asp	285

His	Asn	His	Asp	Thr	Gly	Glu	Gly	Gly	Gln	Gly	Gly	Glu	Glu	Pro
				290					295					300
Arg	Gly	Gly	Arg	Trp	Asp	Glu	Asp	Glu	Gln	Trp	Ser	Val	Val	Val
				305					310					315
Glu	Cys	Glu	Asp	Cys	Glu	Leu	Ile	Pro	Ala	Asp	His	Val	Ile	Val
				320					325					330
Thr	Val	Ser	Leu	Gly	Val	Leu	Lys	Arg	Gln	Tyr	Thr	Ser	Phe	Phe
				335					340					345
Arg	Pro	Gly	Leu	Pro	Thr	Glu	Lys	Val	Ala	Ala	Ile	His	Arg	Leu
				350					355					360
Gly	Ile	Gly	Thr	Thr	Asp	Lys	Ile	Phe	Leu	Glu	Phe	Glu	Glu	Pro
				365					370					375
Phe	Trp	Gly	Pro	Glu	Cys	Asn	Ser	Leu	Gln	Phe	Val	Trp	Glu	Asp
				380					385					390
Glu	Ala	Glu	Ser	His	Thr	Leu	Thr	Tyr	Pro	Pro	Glu	Leu	Trp	Tyr
				395					400					405
Arg	Lys	Ile	Cys	Gly	Phe	Asp	Val	Leu	Tyr	Pro	Pro	Glu	Arg	Tyr
				410					415					420
Gly	His	Val	Leu	Ser	Gly	Trp	Ile	Cys	Gly	Glu	Glu	Ala	Leu	Val
				425					430					435
Met	Glu	Lys	Cys	Asp	Asp	Glu	Ala	Val	Ala	Glu	Ile	Cys	Thr	Glu
				440					445					450
Met	Leu	Arg	Gln	Phe	Thr	Gly	Asn	Pro	Asn	Ile	Pro	Lys	Pro	Arg
				455					460					465
Arg	Ile	Leu	Arg	Ser	Ala	Trp	Gly	Ser	Asn	Pro	Tyr	Phe	Arg	Gly
				470					475					480
Ser	Tyr	Ser	Tyr	Thr	Gln	Val	Gly	Ser	Ser	Gly	Ala	Asp	Val	Glu
				485					490					495
Lys	Leu	Ala	Lys	Pro	Leu	Pro	Tyr	Thr	Glu	Ser	Ser	Lys	Thr	Ala
				500					505					510
Pro	Met	Gln	Val	Leu	Phe	Ser	Gly	Glu	Ala	Thr	His	Arg	Lys	Tyr
				515					520					525
Tyr	Ser	Thr	Thr	His	Gly	Ala	Leu	Leu	Ser	Gly	Gln	Arg	Glu	Ala
				530					535					540
Ala	Arg	Leu	Ile	Glu	Met	Tyr	Arg	Asp	Leu	Phe	Gln	Gln	Gly	Thr
				545					550					555

<210> 4

<211> 337

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1302741CD1

<400> 4

Met	Ala	Leu	Gln	Thr	Leu	Gln	Ser	Ser	Trp	Val	Thr	Phe	Arg	Lys
1				5					10					15
Ile	Leu	Ser	His	Phe	Pro	Glu	Glu	Leu	Ser	Leu	Ala	Phe	Val	Tyr
				20					25					30
Gly	Ser	Gly	Val	Tyr	Arg	Gln	Ala	Gly	Pro	Ser	Ser	Asp	Gln	Lys
				35					40					45
Asn	Ala	Met	Leu	Asp	Phe	Val	Phe	Thr	Val	Asp	Asp	Pro	Val	Ala
				50					55					60
Trp	His	Ser	Lys	Asn	Leu	Lys	Lys	Asn	Trp	Ser	His	Tyr	Ser	Phe
				65					70					75
Leu	Lys	Val	Leu	Gly	Pro	Lys	Ile	Ile	Thr	Ser	Ile	Gln	Asn	Asn
				80					85					90
Tyr	Gly	Ala	Gly	Val	Tyr	Tyr	Asn	Ser	Leu	Ile	Met	Cys	Asn	Gly
				95					100					105
Arg	Leu	Ile	Lys	Tyr	Gly	Val	Ile	Ser	Thr	Asn	Val	Leu	Ile	Glu
				110					115					120
Asp	Leu	Leu	Asn	Trp	Asn	Asn	Leu	Tyr	Ile	Ala	Gly	Arg	Leu	Gln

Lys	Pro	Val	Lys	125	Ile	Ile	Ser	Val	Asn	130	Glu	Asp	Val	Thr	Leu	135	Arg
Ser	Ala	Leu	Asp	140	Arg	Asn	Leu	Lys	Ser	145	Ala	Val	Thr	Ala	Ala	150	Phe
Leu	Met	Leu	Pro	155	Glu	Ser	Phe	Ser	Glu	160	Glu	Asp	Leu	Phe	Ile	165	Glu
Ile	Ala	Gly	Leu	170	Ser	Tyr	Ser	Gly	Asp	175	Phe	Arg	Met	Val	Val	180	Gly
Glu	Asp	Lys	Thr	185	Lys	Val	Leu	Asn	Ile	190	Val	Lys	Pro	Asn	Ile	195	Ala
His	Phe	Arg	Glu	200	Leu	Tyr	Gly	Ser	Ile	205	Leu	Gln	Glu	Asn	Pro	210	Gln
Val	Val	Tyr	Lys	215	Ser	Gln	Gln	Gly	Trp	220	Leu	Glu	Ile	Asp	Lys	225	Ser
Pro	Glu	Gly	Gln	230	Phe	Thr	Gln	Leu	Met	235	Thr	Leu	Pro	Lys	Thr	240	Leu
Gln	Gln	Gln	Ile	245	Asn	His	Ile	Met	Asp	250	Pro	Pro	Gly	Lys	Asn	255	Arg
Asp	Val	Glu	Glu	260	Thr	Leu	Phe	Gln	Val	265	Ala	His	Asp	Pro	Asp	270	Cys
Gly	Asp	Val	Val	275	Arg	Leu	Gly	Leu	Ser	280	Ala	Ile	Val	Arg	Pro	285	Ser
Ser	Ile	Arg	Gln	290	Ser	Thr	Lys	Gly	Ile	295	Phe	Thr	Ala	Gly	Leu	300	Lys
Lys	Ser	Val	Ile	305	Tyr	Ser	Ser	Leu	Lys	310	Leu	His	Lys	Met	Trp	315	Lys
Gly	Trp	Leu	Arg	320	Lys	Thr	Ser			325						330	
				335													

<210> 5
 <211> 109
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1541028CD1

Met	Ser	Ala	Asn	Thr	Phe	Gly	Asn	Ala	Gly	Phe	Ser	Val	Leu	Leu
1				5					10					15
Pro	Gly	Ala	Arg	Leu	Glu	Gly	Arg	Cys	Gly	Pro	Thr	Asn	Ala	Arg
				20					25					30
Val	Arg	Cys	His	Leu	Gly	Leu	Lys	Ile	Pro	Pro	Gly	Cys	Glu	Leu
				35					40					45
Val	Val	Gly	Gly	Glu	Pro	Gln	Cys	Trp	Ala	Glu	Gly	His	Cys	Leu
				50					55					60
Leu	Val	Asp	Asp	Ser	Phe	Leu	His	Thr	Val	Ala	His	Asn	Gly	Ser
				65					70					75
Pro	Glu	Asp	Gly	Pro	Arg	Val	Val	Phe	Ile	Val	Asp	Leu	Trp	His
				80					85					90
Pro	Asn	Val	Ala	Gly	Ala	Glu	Arg	Gln	Ala	Leu	Asp	Phe	Val	Phe
				95					100					105
Ala	Pro	Asp	Pro											

<210> 6
 <211> 385
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature

<223> Incyte ID No: 1597687CD1

<400> 6

Met	Lys	Met	Leu	Pro	Gly	Val	Gly	Val	Phe	Gly	Thr	Gly	Ser	Ser	
1				5					10					15	
Ala	Arg	Val	Leu	Val	Pro	Leu	Leu	Arg	Ala	Glu	Gly	Phe	Thr	Val	
				20					25					30	
Glu	Ala	Leu	Trp	Gly	Lys	Thr	Glu	Glu	Glu	Ala	Lys	Gln	Leu	Ala	
				35					40					45	
Glu	Glu	Met	Asn	Ile	Ala	Phe	Tyr	Thr	Ser	Arg	Thr	Asp	Asp	Ile	
				50					55					60	
Leu	Leu	His	Gln	Asp	Val	Asp	Leu	Val	Cys	Ile	Ser	Ile	Pro	Pro	
				65					70					75	
Pro	Leu	Thr	Arg	Gln	Ile	Ser	Val	Lys	Ala	Leu	Gly	Ile	Gly	Lys	
				80					85					90	
Asn	Val	Val	Cys	Glu	Lys	Ala	Ala	Thr	Ser	Val	Asp	Ala	Phe	Arg	
				95					100					105	
Met	Val	Thr	Ala	Ser	Arg	Tyr	Tyr	Pro	Gln	Leu	Met	Ser	Leu	Val	
				110					115					120	
Gly	Asn	Val	Leu	Arg	Phe	Leu	Pro	Ala	Phe	Val	Arg	Met	Lys	Gln	
				125					130					135	
Leu	Ile	Ser	Glu	His	Tyr	Val	Gly	Ala	Val	Met	Ile	Cys	Asp	Ala	
				140					145					150	
Arg	Ile	Tyr	Ser	Gly	Ser	Leu	Leu	Ser	Pro	Ser	Tyr	Gly	Trp	Ile	
				155					160					165	
Cys	Asp	Glu	Leu	Met	Gly	Gly	Gly	Gly	Leu	His	Thr	Met	Gly	Thr	
				170					175					180	
Tyr	Ile	Val	Asp	Leu	Leu	Thr	His	Leu	Thr	Gly	Arg	Arg	Ala	Glu	
				185					190					195	
Lys	Val	His	Gly	Leu	Leu	Lys	Thr	Phe	Val	Arg	Gln	Asn	Ala	Ala	
				200					205					210	
Ile	Arg	Gly	Ile	Arg	His	Val	Thr	Ser	Asp	Asp	Phe	Cys	Phe	Phe	
				215					220					225	
Gln	Met	Leu	Met	Gly	Gly	Gly	Val	Cys	Ser	Thr	Val	Thr	Leu	Asn	
				230					235					240	
Phe	Asn	Met	Pro	Gly	Ala	Phe	Val	His	Glu	Val	Met	Val	Val	Gly	
				245					250					255	
Ser	Ala	Gly	Arg	Leu	Val	Ala	Arg	Gly	Ala	Asp	Leu	Tyr	Gly	Gln	
				260					265					270	
Lys	Asn	Ser	Ala	Thr	Gln	Glu	Glu	Leu	Leu	Leu	Arg	Asp	Ser	Leu	
				275					280					285	
Ala	Val	Gly	Ala	Gly	Leu	Pro	Glu	Gln	Gly	Pro	Gln	Asp	Val	Pro	
				290					295					300	
Leu	Leu	Tyr	Leu	Lys	Gly	Met	Val	Tyr	Met	Val	Gln	Ala	Leu	Arg	
				305					310					315	
Gln	Ser	Phe	Gln	Gly	Gln	Gly	Asp	Arg	Arg	Thr	Trp	Asp	Arg	Thr	
				320					325					330	
Pro	Val	Ser	Met	Ala	Ala	Ser	Phe	Glu	Asp	Gly	Leu	Tyr	Met	Gln	
				335					340					345	
Ser	Val	Val	Asp	Ala	Ile	Lys	Arg	Ser	Ser	Arg	Ser	Gly	Glu	Trp	
				350					355					360	
Glu	Ala	Val	Glu	Val	Leu	Thr	Glu	Glu	Pro	Asp	Thr	Asn	Gln	Asn	
				365					370					375	
Leu	Cys	Glu	Ala	Leu	Gln	Arg	Asn	Asn	Leu						
				380					385						

<210> 7

<211> 312

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1690348CD1

<400> 7

Met	Ala	Cys	Ala	Ala	Val	Met	Ile	Pro	Gly	Leu	Leu	Arg	Cys	Ser	1	5	10	15
Val	Gly	Ala	Ile	Arg	Ile	Glu	Ala	Ala	Ser	Leu	Arg	Leu	Thr	Leu	20	25	30	35
Ser	Thr	Leu	Arg	His	Leu	Thr	Leu	Thr	Ser	Ile	Met	Lys	Ser	Lys	40	45	50	55
Arg	Lys	Thr	Asp	His	Met	Glu	Arg	Thr	Ala	Ser	Val	Leu	Arg	Arg	60	65	70	75
Glu	Ile	Val	Ala	Ala	Ala	Lys	Val	Cys	Gly	Ala	Ala	Ser	Glu	Ser	80	85	90	95
Pro	Ser	Val	Lys	Ser	Leu	Arg	Leu	Leu	Val	Ala	Asp	Gln	Asp	Phe	100	105	110	115
Ser	Phe	Lys	Ala	Gly	Gln	Trp	Val	Asp	Phe	Phe	Ile	Pro	Gly	Val	120	125	130	135
Ser	Val	Val	Gly	Gly	Phe	Ser	Ile	Cys	Ser	Ser	Pro	Arg	Leu	Leu	140	145	150	155
Glu	Gln	Glu	Arg	Val	Ile	Glu	Leu	Ala	Val	Lys	Tyr	Thr	Asn	His	160	165	170	175
Pro	Pro	Ala	Leu	Trp	Val	His	Asn	Thr	Cys	Thr	Leu	Asp	Cys	Glu	180	185	190	195
Val	Ala	Val	Arg	Val	Gly	Gly	Glu	Phe	Phe	Phe	Asp	Pro	Gln	Pro	200	205	210	215
Ala	Asp	Ala	Ser	Arg	Asn	Leu	Val	Leu	Ile	Ala	Gly	Gly	Val	Gly	220	225	230	235
Ile	Asn	Pro	Leu	Leu	Ser	Ile	Leu	Arg	His	Ala	Ala	Asp	Leu	Leu	240	245	250	255
Arg	Glu	Gln	Ala	Asn	Lys	Arg	Asn	Gly	Tyr	Glu	Ile	Gly	Thr	Ile	260	265	270	275
Lys	Leu	Phe	Tyr	Ser	Ala	Lys	Asn	Thr	Ser	Glu	Leu	Leu	Phe	Lys	280	285	290	295
Lys	Asn	Ile	Leu	Asp	Leu	Val	Asn	Glu	Phe	Pro	Glu	Lys	Ile	Ala	300	305	310	
Cys	Ser	Leu	His	Val	Thr	Lys	Gln	Thr	Thr	Gln	Ile	Asn	Ala	Glu				
Leu	Lys	Pro	Tyr	Ile	Thr	Glu	Gly	Arg	Ile	Thr	Glu	Lys	Glu	Ile				
Arg	Asp	His	Ile	Ser	Lys	Glu	Thr	Leu	Phe	Tyr	Ile	Cys	Gly	Pro				
Pro	Pro	Met	Thr	Asp	Phe	Phe	Ser	Lys	Gln	Leu	Glu	Asn	Asn	His				
Val	Pro	Lys	Glu	His	Ile	Cys	Phe	Glu	Lys	Trp	Trp							

<210> 8

<211> 160

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1865603CD1

<400> 8

Met	Leu	Arg	His	Leu	Pro	Ser	Arg	Leu	Pro	Val	Lys	Met	Trp	Gly	1	5	10	15
Arg	Thr	Leu	Glu	Lys	Gln	Ser	Trp	Arg	Asp	Ser	Ser	Gln	Thr	Pro	20	25	30	35
Pro	Pro	Cys	Leu	Ile	Arg	Arg	Leu	Asp	His	Ile	Val	Met	Thr	Val	40	45	50	55
Lys	Ser	Ile	Lys	Asp	Thr	Thr	Met	Phe	Tyr	Ser	Lys	Ile	Leu	Gly	60	65	70	75
Met	Glu	Val	Met	Thr	Phe	Lys	Glu	Asp	Arg	Lys	Ala	Leu	Cys	Phe				

Gly	Asp	Gln	Lys	Phe	Asn	Leu	His	Glu	Val	Gly	Lys	Glu	Phe	Glu	
				80					85					90	
Pro	Lys	Ala	Ala	His	Pro	Val	Pro	Gly	Ser	Leu	Asp	Ile	Cys	Leu	
				95					100					105	
Ile	Thr	Glu	Val	Pro	Leu	Glu	Glu	Met	Ile	Gln	His	Leu	Lys	Ala	
				110					115					120	
Cys	Asp	Val	Pro	Ile	Glu	Glu	Gly	Pro	Val	Pro	Arg	Thr	Gly	Ala	
				125					130					135	
Lys	Gly	Pro	Ile	Met	Ser	Ile	Tyr	Phe	Arg	Asp	Pro	Asp	Arg	Asn	
				140					145					150	
Leu	Ile	Glu	Val	Ser	Asn	Tyr	Ile	Ser	Ser						
				155					160						

<210> 9
 <211> 487
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1976472CD1

<400> 9

Met	Asp	Trp	Ile	Arg	Leu	Thr	Lys	Ser	Gly	Lys	Asp	Leu	Thr	Gly	
1				5					10					15	
Leu	Lys	Gly	Arg	Leu	Ile	Glu	Val	Thr	Glu	Glu	Glu	Leu	Lys	Lys	
				20					25					30	
His	Asn	Lys	Lys	Asp	Asp	Cys	Trp	Ile	Cys	Ile	Arg	Gly	Phe	Val	
				35					40					45	
Tyr	Asn	Val	Ser	Pro	Tyr	Met	Glu	Tyr	His	Pro	Gly	Gly	Glu	Asp	
				50					55					60	
Glu	Leu	Met	Arg	Ala	Ala	Gly	Ser	Asp	Gly	Thr	Glu	Leu	Phe	Asp	
				65					70					75	
Gln	Val	His	Arg	Trp	Val	Asn	Tyr	Glu	Ser	Met	Leu	Lys	Glu	Cys	
				80					85					90	
Leu	Val	Gly	Arg	Met	Ala	Ile	Lys	Pro	Ala	Val	Leu	Lys	Asp	Tyr	
				95					100					105	
Arg	Glu	Glu	Glu	Lys	Lys	Val	Leu	Asn	Gly	Met	Leu	Pro	Lys	Ser	
				110					115					120	
Gln	Val	Thr	Asp	Thr	Leu	Ala	Lys	Glu	Gly	Pro	Ser	Tyr	Pro	Ser	
				125					130					135	
Tyr	Asp	Trp	Phe	Gln	Thr	Asp	Ser	Leu	Val	Thr	Ile	Ala	Ile	Tyr	
				140					145					150	
Thr	Lys	Gln	Lys	Asp	Ile	Asn	Leu	Asp	Ser	Ile	Ile	Val	Asp	His	
				155					160					165	
Gln	Asn	Asp	Ser	Phe	Arg	Ala	Glu	Thr	Ile	Ile	Lys	Asp	Cys	Leu	
				170					175					180	
Tyr	Leu	Ile	His	Ile	Gly	Leu	Ser	His	Glu	Val	Gln	Glu	Asp	Phe	
				185					190					195	
Ser	Val	Arg	Val	Val	Glu	Ser	Val	Gly	Lys	Ile	Glu	Ile	Val	Leu	
				200					205					210	
Gln	Lys	Lys	Glu	Asn	Thr	Ser	Trp	Asp	Phe	Leu	Gly	His	Pro	Leu	
				215					220					225	
Lys	Asn	His	Asn	Ser	Leu	Ile	Pro	Arg	Lys	Asp	Thr	Gly	Leu	Tyr	
				230					235					240	
Tyr	Arg	Lys	Cys	Gln	Leu	Ile	Ser	Lys	Glu	Asp	Val	Thr	His	Asp	
				245					250					255	
Thr	Arg	Leu	Phe	Cys	Leu	Met	Leu	Pro	Pro	Ser	Thr	His	Leu	Gln	
				260					265					270	
Val	Pro	Ile	Gly	Gln	His	Val	Tyr	Leu	Lys	Leu	Pro	Ile	Thr	Gly	
				275					280					285	
Thr	Glu	Ile	Val	Lys	Pro	Tyr	Thr	Pro	Val	Ser	Gly	Ser	Leu	Leu	
				290					295					300	
Ser	Glu	Phe	Lys	Glu	Pro	Val	Leu	Pro	Asn	Asn	Lys	Tyr	Ile	Tyr	

Phe	Leu	Ile	Lys	305	Ile	Tyr	Pro	Thr	Gly	310	Leu	Phe	Thr	Pro	Glu	315
Asp	Arg	Leu	Gln	320	Gly	Asp	Phe	Val	Ser	325	Val	Ser	Ser	Pro	Glu	330
Gly	Asn	Phe	Lys	335	Ile	Ser	Lys	Phe	Gln	340	Leu	Glu	Asp	Leu	Phe	345
Leu	Leu	Ala	Ala	350	Gly	Thr	Gly	Phe	Thr	355	Pro	Met	Val	Lys	Ile	360
Asn	Tyr	Ala	Leu	365	Thr	Asp	Ile	Pro	Ser	370	Leu	Arg	Lys	Val	Lys	375
Met	Phe	Phe	Asn	380	Lys	Thr	Glu	Asp	Asp	385	Ile	Ile	Trp	Arg	Ser	390
Leu	Glu	Lys	Leu	395	Ala	Phe	Lys	Asp	Lys	400	Leu	Asp	Val	Glu	Phe	405
Val	Leu	Ser	Ala	410	Pro	Ile	Ser	Glu	Trp	415	Asn	Gly	Lys	Gln	Gly	420
Ile	Ser	Pro	Ala	425	Leu	Leu	Ser	Glu	Phe	430	Leu	Lys	Arg	Asn	Leu	435
Lys	Ser	Lys	Val	440	Leu	Val	Cys	Ile	Cys	445	Gly	Pro	Val	Pro	Phe	450
Glu	Gln	Gly	Val	455	Arg	Leu	Leu	His	Asp	460	Leu	Asn	Phe	Ser	Lys	465
Glu	Ile	His	Ser	470	Phe	Thr	Ala			475						480
				485												

<210> 10

<211> 524

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2050821CD1

<400> 10

Met	Ser	Leu	Leu	Ser	Leu	Pro	Trp	Leu	Gly	Leu	Arg	Pro	Val	Ala		
1				5					10					15		
Met	Ser	Pro	Trp	Leu	Leu	Leu	Leu	Leu	Val	Val	Gly	Ser	Trp	Leu		
				20					25					30		
Leu	Ala	Arg	Ile	Leu	Ala	Trp	Thr	Tyr	Ala	Phe	Tyr	Asn	Asn	Cys		
				35					40					45		
Arg	Arg	Leu	Gln	Cys	Phe	Pro	Gln	Pro	Pro	Lys	Arg	Asn	Trp	Phe		
				50					55					60		
Trp	Gly	His	Leu	Gly	Leu	Ile	Thr	Pro	Thr	Glu	Glu	Gly	Leu	Lys		
				65					70					75		
Asp	Ser	Thr	Gln	Met	Ser	Ala	Thr	Tyr	Ser	Gln	Gly	Phe	Thr	Val		
				80					85					90		
Trp	Leu	Gly	Pro	Ile	Ile	Pro	Phe	Ile	Val	Leu	Cys	His	Pro	Asp		
				95					100					105		
Thr	Ile	Arg	Ser	Ile	Thr	Asn	Ala	Ser	Ala	Ala	Ile	Ala	Pro	Lys		
				110					115					120		
Asp	Asn	Leu	Phe	Ile	Arg	Phe	Leu	Lys	Pro	Trp	Leu	Gly	Glu	Gly		
				125					130					135		
Ile	Leu	Leu	Ser	Gly	Gly	Asp	Lys	Trp	Ser	Arg	His	Arg	Arg	Met		
				140					145					150		
Leu	Thr	Pro	Ala	Phe	His	Phe	Asn	Ile	Leu	Lys	Ser	Tyr	Ile	Thr		
				155					160					165		
Ile	Phe	Asn	Lys	Ser	Ala	Asn	Ile	Met	Leu	Asp	Lys	Trp	Gln	His		
				170					175					180		
Leu	Ala	Ser	Glu	Gly	Ser	Ser	Arg	Leu	Asp	Met	Phe	Glu	His	Ile		
				185					190					195		
Ser	Leu	Met	Thr	Leu	Asp	Ser	Leu	Gln	Lys	Cys	Ile	Phe	Ser	Phe		
				200					205					210		

Asp	Ser	His	Cys	Gln	Glu	Arg	Pro	Ser	Glu	Tyr	Ile	Ala	Thr	Ile	
				215					220					225	
Leu	Glu	Leu	Ser	Ala	Leu	Val	Glu	Lys	Arg	Ser	Gln	His	Ile	Leu	
				230					235					240	
Gln	His	Met	Asp	Phe	Leu	Tyr	Tyr	Leu	Ser	His	Asp	Gly	Arg	Arg	
				245					250					255	
Phe	His	Arg	Ala	Cys	Arg	Leu	Val	His	Asp	Phe	Thr	Asp	Ala	Val	
				260					265					270	
Ile	Arg	Glu	Arg	Arg	Arg	Thr	Leu	Pro	Thr	Gln	Gly	Ile	Asp	Asp	
				275					280					285	
Phe	Phe	Lys	Asp	Lys	Ala	Lys	Ser	Lys	Thr	Leu	Asp	Phe	Ile	Asp	
				290					295					300	
Val	Leu	Leu	Leu	Ser	Lys	Asp	Glu	Asp	Gly	Lys	Ala	Leu	Ser	Asp	
				305					310					315	
Glu	Asp	Ile	Arg	Ala	Glu	Ala	Asp	Thr	Phe	Met	Phe	Gly	Gly	His	
				320					325					330	
Asp	Thr	Thr	Ala	Ser	Gly	Leu	Ser	Trp	Val	Leu	Tyr	Asn	Leu	Ala	
				335					340					345	
Arg	His	Pro	Glu	Tyr	Gln	Glu	Arg	Cys	Arg	Gln	Glu	Val	Gln	Glu	
				350					355					360	
Leu	Leu	Lys	Asp	Arg	Asp	Pro	Lys	Glu	Ile	Glu	Trp	Asp	Asp	Leu	
				365					370					375	
Ala	Gln	Leu	Pro	Phe	Leu	Thr	Met	Cys	Val	Lys	Glu	Ser	Leu	Arg	
				380					385					390	
Leu	His	Pro	Pro	Ala	Pro	Phe	Ile	Ser	Arg	Cys	Cys	Thr	Gln	Asp	
				395					400					405	
Ile	Val	Leu	Pro	Asp	Gly	Arg	Val	Ile	Pro	Lys	Gly	Ile	Thr	Cys	
				410					415					420	
Leu	Ile	Asp	Ile	Ile	Gly	Val	His	His	Asn	Pro	Thr	Val	Trp	Pro	
				425					430					435	
Asp	Pro	Glu	Val	Tyr	Asp	Pro	Phe	Arg	Phe	Asp	Pro	Glu	Asn	Ser	
				440					445					450	
Lys	Gly	Arg	Ser	Pro	Leu	Ala	Phe	Ile	Pro	Phe	Ser	Ala	Gly	Pro	
				455					460					465	
Arg	Asn	Cys	Ile	Gly	Gln	Ala	Phe	Ala	Met	Ala	Glu	Met	Lys	Val	
				470					475					480	
Val	Leu	Ala	Leu	Met	Leu	Leu	His	Phe	Arg	Phe	Leu	Pro	Asp	His	
				485					490					495	
Thr	Glu	Pro	Arg	Arg	Lys	Leu	Glu	Leu	Ile	Met	Arg	Ala	Glu	Gly	
				500					505					510	
Gly	Leu	Trp	Leu	Arg	Val	Glu	Pro	Leu	Asn	Val	Gly	Leu	Gln		
				515					520						

<210> 11
 <211> 144
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2408443CD1

<400> 11
 Met Val Thr Leu Tyr Ser Ser Pro Ser Cys Thr Ser Cys Arg Lys
 1 5 10 15
 Ala Lys Gln Trp Leu Val Asp His Asn Leu Pro Phe Ile Glu Arg
 20 25 30
 Asn Leu Asn Lys Glu Pro Leu Arg Ala Glu Asp Val Lys Ala Met
 35 40 45
 Leu Arg Leu Thr Glu Asp Gly Thr Glu Glu Leu Ile Ser Thr Arg
 50 55 60
 Ser Lys Ile Phe Ser Glu Leu Thr Ile Asp Leu Asp Asp Met Ser
 65 70 75
 Ile Asn Lys Leu Ile Asp Leu Ile Val Met Tyr Pro Ser Leu Leu

	80		85		90
Lys Arg Pro Ile	Ile Leu Asp Asp Gln	Arg Met Gln Ile Gly	Tyr		
	95		100		105
Asn Asp Asp Glu	Ile Arg Arg Phe Leu	Pro Arg Glu Val Arg	Gln		
	110		115		120
Arg Glu Leu Ile	Arg Ala Thr Phe Lys	Ala Asp Phe Ala Glu	Glu		
	125		130		135
Ala Lys Asp Leu	Val Val Glu Glu Gly				
	140				

<210> 12

<211> 373

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2508668CD1

<400> 12

Met Trp Val Cys Ser	Thr Leu Trp Arg Val	Arg Thr Pro Ala Arg	
1	5	10	15
Gln Trp Arg Gly Leu	Leu Pro Ala Ser Gly	Cys His Gly Pro Ala	
	20	25	30
Ala Ser Ser Tyr Ser	Ala Ser Ala Glu Pro	Ala Arg Val Arg Ala	
	35	40	45
Leu Val Tyr Gly His	His Gly Asp Pro Ala	Lys Val Val Glu Leu	
	50	55	60
Lys Asn Leu Glu Leu	Ala Ala Val Arg Gly	Ser Asp Val Arg Val	
	65	70	75
Lys Met Leu Ala Ala	Pro Ile Asn Pro Ser	Asp Ile Asn Met Ile	
	80	85	90
Gln Gly Asn Tyr Gly	Leu Leu Pro Glu Leu	Pro Ala Val Gly Gly	
	95	100	105
Asn Glu Gly Val Ala	Gln Val Val Ala Val	Gly Ser Asn Val Thr	
	110	115	120
Gly Leu Lys Pro Gly	Asp Trp Val Ile Pro	Ala Asn Ala Gly Leu	
	125	130	135
Gly Thr Trp Arg Thr	Glu Ala Val Phe Ser	Glu Glu Ala Leu Ile	
	140	145	150
Gln Val Pro Ser Asp	Ile Pro Leu Gln Ser	Ala Ala Thr Leu Gly	
	155	160	165
Val Asn Pro Cys Thr	Ala Tyr Arg Met Leu	Met Asp Phe Glu Gln	
	170	175	180
Leu Gln Pro Gly Asp	Ser Val Ile Gln Asn	Ala Ser Asn Ser Gly	
	185	190	195
Val Gly Gln Ala Val	Ile Gln Ile Ala Ala	Ala Leu Gly Leu Arg	
	200	205	210
Thr Ile Asn Val Val	Arg Asp Arg Pro Asp	Ile Gln Lys Leu Ser	
	215	220	225
Asp Arg Leu Lys Ser	Leu Gly Ala Glu His	Val Ile Thr Glu Glu	
	230	235	240
Glu Leu Arg Arg Pro	Glu Met Lys Asn Phe	Phe Lys Asp Met Pro	
	245	250	255
Gln Pro Arg Leu Ala	Leu Asn Cys Val Gly	Gly Lys Ser Ser Thr	
	260	265	270
Glu Leu Leu Arg Gln	Leu Ala Arg Gly Gly	Thr Met Val Thr Tyr	
	275	280	285
Gly Gly Met Ala Lys	Gln Pro Val Val Ala	Ser Val Ser Leu Leu	
	290	295	300
Ile Phe Lys Asp Leu	Lys Leu Arg Gly Phe	Trp Leu Ser Gln Trp	
	305	310	315
Lys Lys Asp His Ser	Pro Asp Gln Phe Lys	Glu Leu Ile Leu Thr	
	320	325	330

Leu	Cys	Asp	Leu	Ile	Arg	Arg	Gly	Gln	Leu	Thr	Ala	Pro	Ala	Cys	
				335					340					345	
Ser	Gln	Val	Pro	Leu	Gln	Asp	Tyr	Gln	Ser	Ala	Leu	Glu	Ala	Ser	
				350					355					360	
Met	Lys	Pro	Phe	Ile	Ser	Ser	Lys	Gln	Ile	Leu	Thr	Met			
				365					370						

<210> 13
 <211> 305
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2536830CD1

<400> 13

Met	Ala	Lys	Phe	Val	Ile	Ala	Gly	Arg	Ala	Asp	Cys	Pro	Tyr	Tyr	
1				5					10					15	
Ala	Lys	Thr	Glu	Leu	Val	Ala	Asp	Tyr	Leu	Gln	Lys	Asn	Leu	Pro	
				20					25					30	
Asp	Phe	Arg	Ile	His	Lys	Ile	Thr	Gln	Arg	Pro	Glu	Val	Trp	Glu	
				35					40					45	
Asp	Trp	Leu	Lys	Asp	Val	Cys	Glu	Lys	Asn	Lys	Trp	Ser	His	Lys	
				50					55					60	
Asn	Ser	Pro	Ile	Ile	Trp	Arg	Glu	Leu	Leu	Asp	Arg	Gly	Gly	Lys	
				65					70					75	
Gly	Leu	Leu	Leu	Gly	Gly	Tyr	Asn	Glu	Phe	Leu	Glu	His	Ala	Gln	
				80					85					90	
Leu	Tyr	Tyr	Asp	Val	Thr	Ser	Ser	Met	Thr	Thr	Glu	Leu	Met	Met	
				95					100					105	
Val	Ile	Ala	Gln	Glu	Asn	Leu	Gly	Ala	His	Ile	Glu	Lys	Glu	Gln	
				110					115					120	
Glu	Glu	Glu	Ala	Leu	Lys	Thr	Cys	Ile	Asn	Pro	Leu	Gln	Val	Trp	
				125					130					135	
Ile	Thr	Ser	Ala	Ser	Ala	Pro	Ala	Cys	Tyr	Asn	Leu	Ile	Pro	Ile	
				140					145					150	
Leu	Thr	Ser	Gly	Glu	Val	Phe	Gly	Met	His	Thr	Glu	Ile	Ser	Ile	
				155					160					165	
Thr	Leu	Phe	Asp	Asn	Lys	Gln	Ala	Glu	Glu	His	Leu	Lys	Ser	Leu	
				170					175					180	
Val	Val	Glu	Thr	Gln	Asp	Leu	Ala	Ser	Pro	Val	Leu	Arg	Ser	Val	
				185					190					195	
Ser	Ile	Cys	Thr	Lys	Val	Glu	Glu	Ala	Phe	Arg	Gln	Ala	His	Val	
				200					205					210	
Ile	Val	Val	Leu	Asp	Asp	Ser	Thr	Asn	Lys	Glu	Val	Phe	Thr	Leu	
				215					220					225	
Glu	Asp	Cys	Leu	Arg	Ser	Arg	Val	Pro	Leu	Cys	Arg	Leu	Tyr	Gly	
				230					235					240	
Tyr	Leu	Ile	Glu	Lys	Asn	Ala	His	Glu	Ser	Val	Arg	Val	Ile	Val	
				245					250					255	
Gly	Gly	Arg	Thr	Phe	Val	Asn	Leu	Lys	Thr	Val	Leu	Leu	Met	Arg	
				260					265					270	
Tyr	Ala	Pro	Arg	Ile	Ala	His	Asn	Ile	Ile	Ala	Val	Ala	Leu	Gly	
				275					280					285	
Val	Glu	Gly	Glu	Ala	Lys	Ala	Ile	Leu	Ala	Arg	Lys	Leu	Lys	Thr	
				290					295					300	
Ala	Pro	Ser	Cys	Glu											
				305											

<210> 14

<211> 500
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2645179CD1

<400> 14
 Met Glu Ala Ala Arg Pro Pro Pro Thr Ala Gly Lys Phe Val Val
 1 5 10 15
 Val Gly Gly Gly Ile Ala Gly Val Thr Cys Ala Glu Gln Leu Ala
 20 25 30
 Thr His Phe Pro Ser Glu Asp Ile Leu Leu Val Thr Ala Ser Pro
 35 40 45
 Val Ile Lys Ala Val Thr Asn Phe Lys Gln Ile Ser Lys Ile Leu
 50 55 60
 Glu Glu Phe Asp Val Glu Glu Gln Ser Ser Thr Met Leu Gly Lys
 65 70 75
 Arg Phe Pro Asn Ile Lys Val Ile Glu Ser Gly Val Lys Gln Leu
 80 85 90
 Lys Ser Glu Glu His Cys Ile Val Thr Glu Asp Gly Asn Gln His
 95 100 105
 Val Tyr Lys Lys Leu Cys Leu Cys Ala Gly Ala Lys Pro Lys Leu
 110 115 120
 Ile Cys Glu Gly Asn Pro Tyr Val Leu Gly Ile Arg Asp Thr Asp
 125 130 135
 Ser Ala Gln Glu Phe Gln Lys Gln Leu Thr Lys Ala Lys Arg Ile
 140 145 150
 Met Ile Ile Gly Asn Gly Gly Ile Ala Leu Glu Leu Val Tyr Glu
 155 160 165
 Ile Glu Gly Cys Glu Val Ile Trp Ala Ile Lys Asp Lys Ala Ile
 170 175 180
 Gly Asn Thr Phe Phe Asp Ala Gly Ala Ala Glu Phe Leu Thr Ser
 185 190 195
 Lys Leu Ile Ala Glu Lys Ser Glu Ala Lys Ile Ala His Lys Arg
 200 205 210
 Thr Arg Tyr Thr Thr Glu Gly Arg Lys Lys Glu Ala Arg Ser Lys
 215 220 225
 Ser Lys Ala Asp Asn Val Gly Ser Ala Leu Gly Pro Asp Trp His
 230 235 240
 Glu Gly Leu Asn Leu Lys Gly Thr Lys Glu Phe Ser His Lys Ile
 245 250 255
 His Leu Glu Thr Met Cys Glu Val Lys Lys Ile Tyr Leu Gln Asp
 260 265 270
 Glu Phe Arg Ile Leu Lys Lys Lys Ser Phe Thr Phe Pro Arg Asp
 275 280 285
 His Lys Ser Val Thr Ala Asp Thr Glu Met Trp Pro Val Tyr Val
 290 295 300
 Glu Leu Thr Asn Glu Lys Ile Tyr Gly Cys Asp Phe Ile Val Ser
 305 310 315
 Ala Thr Gly Val Thr Pro Asn Val Glu Pro Phe Leu His Gly Asn
 320 325 330
 Ser Phe Asp Leu Gly Glu Asp Gly Gly Leu Lys Val Asp Asp His
 335 340 345
 Met His Thr Ser Leu Pro Asp Ile Tyr Ala Ala Gly Asp Ile Cys
 350 355 360
 Thr Thr Ser Trp Gln Leu Ser Pro Val Trp Gln Gln Met Arg Leu
 365 370 375
 Trp Thr Gln Ala Arg Gln Met Gly Trp Tyr Ala Ala Lys Cys Met
 380 385 390
 Ala Ala Ala Ser Ser Gly Asp Ser Ile Asp Met Asp Phe Ser Phe
 395 400 405
 Glu Leu Phe Ala His Val Thr Lys Phe Phe Asn Tyr Lys Val Val
 410 415 420
 Leu Leu Gly Lys Tyr Asn Ala Gln Gly Leu Gly Ser Asp His Glu
 425 430 435

Leu Met Leu Arg Cys Thr Lys Gly Arg Glu Tyr Ile Lys Val Val	440	445	450
Met Gln Asn Gly Arg Met Met Gly Ala Val Leu Ile Gly Glu Thr	455	460	465
Asp Leu Glu Glu Thr Phe Glu Asn Leu Ile Leu Asn Gln Met Asn	470	475	480
Leu Ser Ser Tyr Gly Glu Asp Leu Leu Asp Pro Asn Ile Asp Ile	485	490	495
Glu Asp Tyr Phe Asp	500		

<210> 15
 <211> 369
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2754425CD1

<400> 15

Met Val Trp Ala Pro Leu Gly Pro Pro Arg Thr Asp Cys Leu Thr	1	5	10	15
Leu Leu His Thr Pro Ser Lys Asp Ser Pro Lys Met Ser Leu Glu	20	25	30	
Trp Leu Val Ala Trp Ser Trp Ser Leu Asp Gly Leu Arg Asp Cys	35	40	45	
Ile Ala Thr Gly Ile Gln Ser Val Arg Asp Cys Asp Thr Thr Ala	50	55	60	
Val Ile Thr Val Ala Cys Leu Leu Val Leu Phe Val Trp Tyr Cys	65	70	75	
Tyr His Val Gly Arg Glu Gln Pro Arg Pro Tyr Val Ser Val Asn	80	85	90	
Ser Leu Met Gln Ala Ala Asp Ala Asn Gly Leu Gln Asn Gly Tyr	95	100	105	
Val Tyr Cys Gln Ser Pro Glu Cys Val Arg Cys Thr His Asn Glu	110	115	120	
Gly Leu Asn Gln Lys Leu Tyr His Asn Leu Gln Glu Tyr Ala Lys	125	130	135	
Arg Tyr Ser Trp Ser Gly Met Gly Arg Ile His Lys Gly Ile Arg	140	145	150	
Glu Gln Gly Arg Tyr Leu Asn Ser Arg Pro Ser Ile Gln Lys Pro	155	160	165	
Glu Val Phe Phe Leu Pro Asp Leu Pro Thr Thr Pro Tyr Phe Ser	170	175	180	
Arg Asp Ala Gln Lys His Asp Val Glu Val Leu Glu Arg Asn Phe	185	190	195	
Gln Thr Ile Leu Cys Glu Phe Glu Thr Leu Tyr Lys Ala Phe Ser	200	205	210	
Asn Cys Ser Leu Pro Gln Gly Trp Lys Met Asn Ser Thr Pro Ser	215	220	225	
Gly Glu Trp Phe Thr Phe Tyr Leu Val Asn Gln Gly Val Cys Val	230	235	240	
Pro Arg Asn Cys Arg Lys Cys Pro Arg Thr Tyr Arg Leu Leu Gly	245	250	255	
Ser Leu Arg Thr Cys Ile Gly Asn Asn Val Phe Gly Asn Ala Cys	260	265	270	
Ile Ser Val Leu Ser Pro Gly Thr Val Ile Thr Glu His Tyr Gly	275	280	285	
Pro Thr Asn Ile Arg Ile Arg Cys His Leu Gly Leu Lys Thr Pro	290	295	300	
Asn Gly Cys Glu Leu Val Val Gly Gly Glu Pro Gln Cys Trp Ala	305	310	315	
Glu Gly Arg Cys Leu Leu Phe Asp Asp Ser Phe Leu His Ala Ala				

Phe His Glu Gly	320	325	330
Ser Ala Glu Asp Gly	335	Pro Arg Val Val Phe	Met
Val Asp Leu Trp	340	Ala Ala Glu Arg Gln	Ala
His Pro Asn Val Ala	350	355	360
Leu Asp Phe Ile Phe	365	Ala Pro Gly Arg	

<210> 16
 <211> 145
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2821526CD1

<400> 16	
Met Glu Leu Val Gln Val Leu Lys Arg Gly Leu Gln Gln Ile Thr	
1 5 10 15	
Gly His Gly Gly Leu Arg Gly Tyr Leu Arg Val Phe Phe Arg Thr	
20 25 30	
Asn Asp Ala Lys Val Gly Thr Leu Val Gly Glu Asp Lys Tyr Gly	
35 40 45	
Asn Lys Tyr Tyr Glu Asp Asn Lys Gln Phe Phe Gly Arg His Arg	
50 55 60	
Trp Val Val Tyr Thr Thr Glu Met Asn Gly Lys Asn Thr Phe Trp	
65 70 75	
Asp Val Asp Gly Ser Met Val Pro Pro Glu Trp His Arg Trp Leu	
80 85 90	
His Ser Met Thr Asp Asp Pro Pro Thr Thr Lys Pro Leu Thr Ala	
95 100 105	
Arg Lys Phe Ile Trp Thr Asn His Lys Phe Asn Val Thr Gly Thr	
110 115 120	
Pro Glu Gln Tyr Val Pro Tyr Ser Thr Thr Arg Lys Lys Ile Gln	
125 130 135	
Glu Trp Ile Pro Pro Ser Thr Pro Tyr Lys	
140 145	

<210> 17
 <211> 255
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 2876494CD1

<400> 17	
Met Lys Val Leu Ala Thr Ser Phe Val Leu Gly Ser Leu Gly Leu	
1 5 10 15	
Ala Phe Tyr Leu Pro Leu Val Val Thr Thr Pro Lys Thr Leu Ala	
20 25 30	
Ile Pro Glu Lys Leu Gln Glu Ala Val Gly Lys Val Ile Ile Asn	
35 40 45	
Ala Thr Thr Cys Thr Val Thr Cys Gly Leu Gly Tyr Lys Glu Glu	
50 55 60	
Thr Val Cys Glu Val Gly Pro Asp Gly Val Arg Arg Lys Cys Gln	
65 70 75	
Thr Gln Arg Leu Glu Cys Leu Thr Asn Trp Ile Cys Gly Met Leu	
80 85 90	
His Phe Thr Ile Leu Ile Gly Lys Glu Phe Glu Leu Ser Cys Leu	

	95		100		105
Ser Ser Asp Ile	Leu Glu Phe Gly Gln	Glu Ala Phe Arg Phe	Thr		
	110		115		120
Trp Arg Leu Ala	Arg Gly Val Ile Ser	Thr Asp Asp Glu Val	Phe		
	125		130		135
Lys Pro Phe Gln	Ala Asn Ser His Phe	Val Lys Phe Lys Tyr	Ala		
	140		145		150
Gln Glu Tyr Asp	Ser Gly Thr Tyr Arg	Cys Asp Val Gln Leu	Val		
	155		160		165
Lys Asn Leu Arg	Leu Val Lys Arg Leu	Tyr Phe Gly Leu Arg	Val		
	170		175		180
Leu Pro Pro Asn	Leu Val Asn Leu Asn	Phe His Gln Ser Leu	Thr		
	185		190		195
Glu Asp Gln Lys	Leu Ile Asp Glu Gly	Leu Glu Val Asn Leu	Asp		
	200		205		210
Ser Tyr Ser Lys	Pro His His Pro Lys	Trp Lys Lys Lys Val	Ala		
	215		220		225
Ser Ala Leu Gly	Ile Gly Ile Ala Ile	Gly Val Val Gly Gly	Val		
	230		235		240
Leu Val Arg Ile	Val Leu Cys Ala Leu	Arg Gly Gly Leu Gln	Gln		
	245		250		255

<210> 18

<211> 246

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3403225CD1

<400> 18

Met Leu Val Thr	Leu Gly Leu Leu Thr	Ser Phe Phe Ser Phe	Leu
1	5	10	15
Tyr Met Val Ala	Pro Ser Ile Arg Lys	Phe Phe Ala Gly Gly	Val
	20	25	30
Cys Arg Thr Asn	Val Gln Leu Pro Gly	Lys Val Val Val Ile	Thr
	35	40	45
Gly Ala Asn Thr	Gly Ile Gly Lys Glu	Thr Ala Arg Glu Leu	Ala
	50	55	60
Ser Arg Gly Ala	Arg Val Tyr Ile Ala	Cys Arg Asp Val Leu	Lys
	65	70	75
Gly Glu Ser Ala	Ala Ser Glu Ile Arg	Val Asp Thr Lys Asn	Ser
	80	85	90
Gln Val Leu Val	Arg Lys Leu Asp Leu	Ser Asp Thr Lys Ser	Ile
	95	100	105
Arg Ala Phe Ala	Glu Gly Phe Leu Ala	Glu Glu Lys Gln Leu	His
	110	115	120
Ile Leu Ile Asn	Asn Ala Gly Val Met	Met Cys Pro Tyr Ser	Lys
	125	130	135
Thr Ala Asp Gly	Phe Glu Thr His Leu	Gly Val Asn His Leu	Gly
	140	145	150
Thr Gly Val Thr	Thr Tyr Ala Val His	Pro Gly Val Val Arg	Ser
	155	160	165
Glu Leu Val Arg	His Ser Ser Leu Leu	Cys Leu Leu Trp Arg	Leu
	170	175	180
Phe Ser Pro Phe	Val Lys Thr Ala Arg	Glu Gly Ala Gln Thr	Ser
	185	190	195
Leu His Cys Ala	Leu Ala Glu Gly Leu	Glu Pro Leu Ser Gly	Lys
	200	205	210
Tyr Phe Ser Asp	Cys Lys Arg Thr Trp	Val Ser Pro Arg Ala	Arg
	215	220	225
Asn Asn Lys Thr	Ala Glu Arg Leu Trp	Asn Val Ser Cys Glu	Leu
	230	235	240

Leu Gly Ile Arg Trp Glu
245

<210> 19

<211> 467

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 4163943CD1

<400> 19

Met	Ala	Ala	Ala	Ala	Arg	Ala	Arg	Val	Ala	Tyr	Leu	Leu	Arg	Gln	15
1				5					10						15
Leu	Gln	Arg	Ala	Ala	Cys	Gln	Cys	Pro	Thr	His	Ser	His	Thr	Tyr	30
				20					25						30
Ser	Gln	Ala	Pro	Gly	Leu	Ser	Pro	Ser	Gly	Lys	Thr	Thr	Asp	Tyr	45
				35					40						45
Ala	Phe	Glu	Met	Ala	Val	Ser	Asn	Ile	Arg	Tyr	Gly	Ala	Ala	Val	60
				50					55						60
Thr	Lys	Glu	Val	Gly	Met	Asp	Leu	Lys	Asn	Met	Gly	Ala	Lys	Asn	75
				65					70						75
Val	Cys	Leu	Met	Thr	Asp	Lys	Asn	Leu	Ser	Lys	Leu	Pro	Pro	Val	90
				80					85						90
Gln	Val	Ala	Met	Asp	Ser	Leu	Val	Lys	Asn	Gly	Ile	Pro	Phe	Thr	105
				95					100						105
Val	Tyr	Asp	Asn	Val	Arg	Val	Glu	Pro	Thr	Asp	Ser	Ser	Phe	Met	120
				110					115						120
Glu	Ala	Ile	Glu	Phe	Ala	Gln	Lys	Gly	Ala	Phe	Asp	Ala	Tyr	Val	135
				125					130						135
Ala	Val	Gly	Gly	Gly	Ser	Thr	Met	Asp	Thr	Cys	Lys	Ala	Ala	Asn	150
				140					145						150
Leu	Tyr	Ala	Ser	Ser	Pro	His	Ser	Asp	Phe	Leu	Asp	Tyr	Val	Ser	165
				155					160						165
Ala	Pro	Ile	Gly	Lys	Gly	Lys	Pro	Val	Ser	Val	Pro	Leu	Lys	Pro	180
				170					175						180
Leu	Ile	Ala	Val	Pro	Thr	Thr	Ser	Gly	Thr	Gly	Ser	Glu	Thr	Thr	195
				185					190						195
Gly	Val	Ala	Ile	Phe	Asp	Tyr	Glu	His	Leu	Lys	Val	Lys	Ile	Gly	210
				200					205						210
Ile	Thr	Ser	Arg	Ala	Ile	Lys	Pro	Thr	Leu	Gly	Leu	Ile	Asp	Pro	225
				215					220						225
Leu	His	Thr	Leu	His	Met	Pro	Ala	Arg	Val	Val	Ala	Asn	Ser	Gly	240
				230					235						240
Phe	Asp	Val	Leu	Cys	His	Ala	Leu	Glu	Ser	Tyr	Thr	Thr	Leu	Pro	255
				245					250						255
Tyr	His	Leu	Arg	Ser	Pro	Cys	Pro	Ser	Asn	Pro	Ile	Thr	Arg	Pro	270
				260					265						270
Ala	Tyr	Gln	Gly	Ser	Asn	Pro	Ile	Ser	Asp	Ile	Trp	Ala	Ile	His	285
				275					280						285
Ala	Leu	Arg	Ile	Val	Ala	Lys	Tyr	Leu	Lys	Arg	Ala	Val	Arg	Asn	300
				290					295						300
Pro	Asp	Asp	Leu	Glu	Ala	Arg	Ser	His	Met	His	Leu	Ala	Ser	Ala	315
				305					310						315
Phe	Ala	Gly	Ile	Gly	Phe	Gly	Asn	Ala	Gly	Val	His	Leu	Cys	His	330
				320					325						330
Gly	Met	Ser	Tyr	Pro	Ile	Ser	Gly	Leu	Val	Lys	Met	Tyr	Lys	Ala	345
				335					340						345
Lys	Asp	Tyr	Asn	Val	Asp	His	Pro	Leu	Val	Pro	His	Gly	Leu	Ser	360
				350					355						360
Val	Val	Leu	Thr	Ser	Pro	Ala	Val	Phe	Thr	Phe	Thr	Ala	Gln	Met	375
				365					370						375
Phe	Pro	Glu	Arg	His	Leu	Glu	Met	Ala	Glu	Ile	Leu	Gly	Ala	Asp	

Thr Arg Thr Ala	Arg	Ile Gln Asp Ala	Gly Leu Val Leu Ala	Asp	380	385	390
Thr Leu Arg Lys	Phe	Leu Phe Asp Leu	Asp Val Asp Asp Gly	Leu	395	400	405
Ala Ala Val Gly	Tyr	Ser Lys Ala Asp	Ile Pro Ala Leu Val	Lys	410	415	420
Gly Thr Leu Pro	Gln	Glu Arg Val Thr	Lys Leu Ala Pro Arg	Pro	425	430	435
Gln Ser Glu Glu	Asp	Leu Ala Ala Leu	Phe Glu Ala Ser Met	Lys	440	445	450
Leu Tyr					455	460	465

<210> 20
 <211> 317
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 4293484CD1

<400> 20

Met Ala Asp Ser Ala	Gln Ala Gln Lys	Leu Val Tyr Leu Val	Thr
1	5	10	15
Gly Gly Cys Gly Phe	Leu Gly Glu His	Val Val Arg Met Leu	Leu
20	25	30	35
Gln Arg Glu Pro Arg	Leu Gly Glu Leu	Arg Val Phe Asp Gln	His
35	40	45	50
Leu Gly Pro Trp Leu	Glu Glu Leu Lys	Thr Gly Thr Arg Asn	Val
50	55	60	65
Ile Glu Ala Cys Val	Gln Thr Gly Thr	Arg Phe Leu Val Tyr	Thr
65	70	75	80
Ser Ser Met Glu Val	Val Gly Pro Asn	Thr Lys Gly His Pro	Phe
80	85	90	95
Tyr Arg Gly Asn Glu	Asp Thr Pro Tyr	Glu Ala Val His Arg	His
95	100	105	110
Pro Tyr Pro Cys Ser	Lys Ala Leu Ala	Glu Trp Leu Val Leu	Glu
110	115	120	125
Ala Asn Gly Arg Lys	Val Arg Gly Gly	Leu Pro Leu Val Thr	Cys
125	130	135	140
Ala Leu Arg Pro Thr	Gly Ile Tyr Gly	Glu Gly His Gln Ile	Met
140	145	150	155
Arg Asp Phe Tyr Arg	Gln Gly Leu Arg	Leu Gly Gly Trp Leu	Phe
155	160	165	170
Arg Ala Ile Pro Ala	Ser Val Glu His	Gly Arg Val Tyr Val	Gly
170	175	180	185
Asn Val Ala Trp Met	His Val Leu Ala	Ala Arg Glu Leu Glu	Gln
185	190	195	200
Arg Ala Thr Leu Met	Gly Gly Gln Val	Tyr Phe Cys Tyr Asp	Gly
200	205	210	215
Ser Pro Tyr Arg Ser	Tyr Glu Asp Phe	Asn Met Glu Phe Leu	Gly
215	220	225	230
Pro Cys Gly Leu Arg	Leu Val Gly Ala	Arg Pro Leu Leu Pro	Tyr
230	235	240	245
Trp Leu Leu Val Phe	Leu Ala Ala Leu	Asn Ala Leu Leu Gln	Trp
245	250	255	260
Leu Leu Arg Pro Leu	Val Leu Tyr Ala	Pro Leu Leu Asn Pro	Tyr
260	265	270	275
Thr Leu Ala Val Ala	Asn Thr Thr Phe	Thr Val Ser Thr Asp	Lys
275	280	285	290
Ala Gln Arg His Phe	Gly Tyr Glu Pro	Leu Phe Ser Trp Glu	Asp
290	295	300	305
Ser Arg Thr Arg Thr	Ile Leu Trp Val	Gln Ala Ala Thr Gly	Ser

Ala Gln 305 310 315

<210> 21
 <211> 181
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 4440080CD1

<400> 21
 Met Phe Ser Ala Ile Arg Ser Gln His Ser Gly Val Asp Ile Cys
 1 5 10 15
 Ile Asn Asn Ala Gly Leu Ala Arg Pro Asp Thr Leu Leu Ser Gly
 20 25 30
 Ser Thr Ser Gly Trp Lys Asp Met Phe Asn Val Asn Val Leu Ala
 35 40 45
 Leu Ser Ile Cys Thr Arg Glu Ala Tyr Gln Ser Met Lys Glu Arg
 50 55 60
 Asn Val Asp Asp Gly His Ile Ile Asn Ile Asn Ser Met Ser Gly
 65 70 75
 His Arg Val Leu Pro Leu Ser Val Thr His Phe Tyr Ser Ala Thr
 80 85 90
 Lys Tyr Ala Val Thr Ala Leu Thr Glu Gly Leu Arg Gln Glu Leu
 95 100 105
 Arg Glu Ala Gln Thr His Ile Arg Ala Thr Cys Ile Ser Pro Gly
 110 115 120
 Val Val Glu Thr Gln Phe Ala Phe Lys Leu His Asp Lys Asp Pro
 125 130 135
 Glu Lys Ala Ala Ala Thr Tyr Glu Gln Met Lys Cys Leu Lys Pro
 140 145 150
 Glu Asp Val Ala Glu Ala Val Ile Tyr Val Leu Ser Thr Pro Ala
 155 160 165
 His Ile Gln Ile Gly Asp Ile Gln Met Arg Pro Thr Glu Gln Val
 170 175 180
 Thr

<210> 22
 <211> 360
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 5495687CD1

<400> 22
 Met Val Pro Ala Ala Gly Arg Arg Pro Pro Arg Val Met Arg Leu
 1 5 10 15
 Leu Gly Trp Trp Gln Val Leu Leu Trp Val Leu Gly Leu Pro Val
 20 25 30
 Arg Gly Val Glu Val Ala Glu Glu Ser Gly Arg Leu Trp Ser Glu
 35 40 45
 Glu Gln Pro Ala His Pro Leu Gln Val Gly Ala Val Tyr Leu Gly
 50 55 60
 Glu Glu Glu Leu Leu His Asp Pro Met Gly Gln Asp Arg Ala Ala
 65 70 75
 Glu Glu Ala Asn Ala Val Leu Gly Leu Asp Thr Gln Gly Asp His
 80 85 90

Met	Val	Met	Leu	Ser	Val	Ile	Pro	Gly	Glu	Ala	Glu	Asp	Lys	Val	
				95					100					105	
Ser	Ser	Glu	Pro	Ser	Gly	Val	Thr	Cys	Gly	Ala	Gly	Gly	Ala	Glu	
				110					115					120	
Asp	Ser	Arg	Cys	Asn	Val	Arg	Glu	Ser	Leu	Phe	Ser	Leu	Asp	Gly	
				125					130					135	
Ala	Gly	Ala	His	Phe	Pro	Asp	Arg	Glu	Glu	Glu	Tyr	Tyr	Thr	Glu	
				140					145					150	
Pro	Glu	Val	Ala	Glu	Ser	Asp	Ala	Ala	Pro	Thr	Glu	Asp	Ser	Asn	
				155					160					165	
Asn	Thr	Glu	Ser	Leu	Lys	Ser	Pro	Lys	Val	Asn	Cys	Glu	Glu	Arg	
				170					175					180	
Asn	Ile	Thr	Gly	Leu	Glu	Asn	Phe	Thr	Leu	Lys	Ile	Leu	Asn	Met	
				185					190					195	
Ser	Gln	Asp	Leu	Met	Asp	Phe	Leu	Asn	Pro	Asn	Gly	Ser	Asp	Cys	
				200					205					210	
Thr	Leu	Val	Leu	Phe	Tyr	Thr	Pro	Trp	Cys	Arg	Phe	Ser	Ala	Ser	
				215					220					225	
Leu	Ala	Pro	His	Phe	Asn	Ser	Leu	Pro	Arg	Ala	Phe	Pro	Ala	Leu	
				230					235					240	
His	Phe	Leu	Ala	Leu	Asp	Ala	Ser	Gln	His	Ser	Ser	Leu	Ser	Thr	
				245					250					255	
Arg	Phe	Gly	Thr	Val	Ala	Val	Pro	Asn	Ile	Leu	Leu	Phe	Gln	Gly	
				260					265					270	
Ala	Lys	Pro	Met	Ala	Arg	Phe	Asn	His	Thr	Asp	Arg	Thr	Leu	Glu	
				275					280					285	
Thr	Leu	Lys	Ile	Phe	Ile	Phe	Asn	Gln	Thr	Gly	Ile	Glu	Ala	Lys	
				290					295					300	
Lys	Asn	Val	Val	Val	Thr	Gln	Ala	Asp	Gln	Ile	Gly	Pro	Leu	Pro	
				305					310					315	
Ser	Thr	Leu	Ile	Lys	Ser	Val	Asp	Trp	Leu	Leu	Val	Phe	Ser	Leu	
				320					325					330	
Phe	Phe	Leu	Ile	Ser	Phe	Ile	Met	Tyr	Ala	Thr	Ile	Arg	Thr	Glu	
				335					340					345	
Ser	Ile	Arg	Trp	Leu	Ile	Pro	Gly	Gln	Glu	Gln	Glu	His	Val	Glu	
				350					355					360	

<210> 23

<211> 476

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 5527735CD1

<400> 23

Met	Ala	Ser	Pro	Phe	Ser	Gly	Ala	Leu	Gln	Leu	Thr	Asp	Leu	Asp	
1				5					10					15	
Asp	Phe	Ile	Gly	Pro	Ser	Gln	Glu	Cys	Ile	Lys	Pro	Val	Lys	Val	
				20					25					30	
Glu	Lys	Arg	Ala	Gly	Ser	Gly	Val	Ala	Lys	Ile	Arg	Ile	Glu	Asp	
				35					40					45	
Asp	Gly	Ser	Tyr	Phe	Gln	Ile	Asn	Gln	Asp	Gly	Gly	Thr	Arg	Arg	
				50					55					60	
Leu	Glu	Lys	Ala	Lys	Val	Ser	Leu	Asn	Asp	Cys	Leu	Ala	Cys	Ser	
				65					70					75	
Gly	Cys	Ile	Thr	Ser	Ala	Glu	Thr	Val	Leu	Ile	Thr	Gln	Gln	Ser	
				80					85					90	
His	Glu	Glu	Leu	Lys	Lys	Val	Leu	Asp	Ala	Asn	Lys	Met	Ala	Ala	
				95					100					105	
Pro	Ser	Gln	Gln	Arg	Leu	Val	Val	Val	Ser	Val	Ser	Pro	Gln	Ser	
				110					115					120	
Arg	Ala	Ser	Leu	Ala	Ala	Arg	Phe	Gln	Leu	Asn	Pro	Thr	Asp	Thr	

Ala Arg Lys Leu	125	Thr Ser Phe Phe Lys	130	Lys Ile Gly Val His	135
Val Phe Asp Thr	140	Ala Phe Ser Arg His	145	Phe Ser Leu Leu Glu	150
Gln Arg Glu Phe	155	Val Arg Arg Phe Arg	160	Gly Gln Ala Asp Cys	165
Gln Ala Leu Pro	170	Leu Leu Ala Ser Ala	175	Cys Pro Gly Trp Ile	180
Tyr Ala Glu Lys	185	Thr His Gly Ser Phe	190	Ile Leu Pro His Ile	195
Thr Ala Arg Ser	200	Pro Gln Gln Val Met	205	Gly Ser Leu Val Lys	210
Phe Phe Ala Gln	215	Gln Gln His Leu Thr	220	Pro Asp Lys Ile Tyr	225
Val Thr Val Met	230	Pro Cys Tyr Asp Lys	235	Lys Leu Glu Ala Ser	240
Pro Asp Phe Phe	245	Asn Gln Glu His Gln	250	Thr Arg Asp Val Asp	255
Val Leu Thr Thr	260	Gly Glu Val Phe Arg	265	Leu Leu Glu Glu Glu	270
Val Ser Leu Pro	275	Asp Leu Glu Pro Ala	280	Pro Leu Asp Ser Leu	285
Ser Gly Ala Ser	290	Ala Glu Glu Pro Thr	295	Ser His Arg Gly Gly	300
Ser Gly Gly Tyr	305	Leu Glu His Val Phe	310	Arg His Ala Ala Arg	315
Leu Phe Gly Ile	320	His Val Ala Glu Val	325	Thr Tyr Lys Pro Leu	330
Asn Lys Asp Phe	335	Gln Glu Val Thr Leu	340	Glu Lys Glu Gly Gln	345
Leu Leu His Phe	350	Ala Met Ala Tyr Gly	355	Phe Arg Asn Ile Gln	360
Leu Val Gln Arg	365	Leu Lys Arg Gly Arg	370	Cys Pro Tyr His Tyr	375
Glu Val Met Ala	380	Cys Pro Ser Gly Cys	385	Leu Asn Gly Gly Gly	390
Leu Gln Ala Pro	395	Asp Arg Pro Ser Arg	400	Glu Leu Leu Gln His	405
Glu Arg Leu Tyr	410	Gly Met Val Arg Ala	415	Glu Ala Pro Glu Asp	420
Pro Gly Val Gln	425	Glu Leu Tyr Thr His	430	Trp Leu Gln Gly Thr	435
Ser Glu Cys Ala	440	Gly Arg Leu Leu His	445	Thr Gln Tyr His Ala	450
Glu Lys Ala Ser	455	Thr Gly Leu Gly Ile	460	Arg Trp	465
	470		475		

<210> 24

<211> 621

<212> PRT

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 5540437CD1

<400> 24

Met Ser Gly Cys Gly	1	Leu Phe Leu Arg Thr	5	Thr Ala Ala Ala Arg	15
Ala Cys Arg Gly Leu	20	Val Val Ser Thr Ala	25	Asn Arg Arg Leu Leu	30
Arg Thr Ser Pro Pro	35	Val Arg Ala Phe Ala	40	Lys Glu Leu Phe Leu	45

Gly	Lys	Ile	Lys	Lys	Lys	Glu	Val	Phe	Pro	Phe	Pro	Glu	Val	Ser
				50					55					60
Gln	Asp	Glu	Leu	Asn	Glu	Ile	Asn	Gln	Phe	Leu	Gly	Pro	Val	Glu
				65					70					75
Lys	Phe	Phe	Thr	Glu	Glu	Val	Asp	Ser	Arg	Lys	Ile	Asp	Gln	Glu
				80					85					90
Gly	Lys	Ile	Pro	Asp	Glu	Thr	Leu	Glu	Lys	Leu	Lys	Ser	Leu	Gly
				95					100					105
Leu	Phe	Gly	Leu	Gln	Val	Pro	Glu	Glu	Tyr	Gly	Gly	Leu	Gly	Phe
				110					115					120
Ser	Asn	Thr	Met	Tyr	Ser	Arg	Leu	Gly	Glu	Ile	Ile	Ser	Met	Asp
				125					130					135
Gly	Ser	Ile	Thr	Val	Thr	Leu	Ala	Ala	His	Gln	Ala	Ile	Gly	Leu
				140					145					150
Lys	Gly	Ile	Ile	Leu	Ala	Gly	Thr	Glu	Glu	Gln	Lys	Ala	Lys	Tyr
				155					160					165
Leu	Pro	Lys	Leu	Ala	Ser	Gly	Glu	His	Ile	Ala	Ala	Phe	Cys	Leu
				170					175					180
Thr	Glu	Pro	Ala	Ser	Gly	Ser	Asp	Ala	Ala	Ser	Ile	Arg	Ser	Arg
				185					190					195
Ala	Thr	Leu	Ser	Glu	Asp	Lys	Lys	His	Tyr	Ile	Leu	Asn	Gly	Ser
				200					205					210
Lys	Val	Trp	Ile	Thr	Asn	Gly	Gly	Leu	Ala	Asn	Ile	Phe	Thr	Val
				215					220					225
Phe	Ala	Lys	Thr	Glu	Val	Val	Asp	Ser	Asp	Gly	Ser	Val	Lys	Asp
				230					235					240
Lys	Ile	Thr	Ala	Phe	Ile	Val	Glu	Arg	Asp	Phe	Gly	Gly	Val	Thr
				245					250					255
Asn	Gly	Lys	Pro	Glu	Asp	Lys	Leu	Gly	Ile	Arg	Gly	Ser	Asn	Thr
				260					265					270
Cys	Glu	Val	His	Phe	Glu	Asn	Thr	Lys	Ile	Pro	Val	Glu	Asn	Ile
				275					280					285
Leu	Gly	Glu	Val	Gly	Asp	Gly	Phe	Lys	Val	Ala	Met	Asn	Ile	Leu
				290					295					300
Asn	Ser	Gly	Arg	Phe	Ser	Met	Gly	Ser	Val	Val	Ala	Gly	Leu	Leu
				305					310					315
Lys	Arg	Leu	Ile	Glu	Met	Thr	Ala	Glu	Tyr	Ala	Cys	Thr	Arg	Lys
				320					325					330
Gln	Phe	Asn	Lys	Arg	Leu	Ser	Glu	Phe	Gly	Leu	Ile	Gln	Glu	Lys
				335					340					345
Phe	Ala	Leu	Met	Ala	Gln	Lys	Ala	Tyr	Val	Met	Glu	Ser	Met	Thr
				350					355					360
Tyr	Leu	Thr	Ala	Gly	Met	Leu	Asp	Gln	Pro	Gly	Phe	Pro	Asp	Cys
				365					370					375
Ser	Ile	Glu	Ala	Ala	Met	Val	Lys	Val	Phe	Ser	Ser	Glu	Ala	Ala
				380					385					390
Trp	Gln	Cys	Val	Ser	Glu	Ala	Leu	Gln	Ile	Leu	Gly	Gly	Leu	Gly
				395					400					405
Tyr	Thr	Arg	Asp	Tyr	Pro	Tyr	Glu	Arg	Ile	Leu	Arg	Asp	Thr	Arg
				410					415					420
Ile	Leu	Leu	Ile	Phe	Glu	Gly	Thr	Asn	Glu	Ile	Leu	Arg	Met	Tyr
				425					430					435
Ile	Ala	Leu	Thr	Gly	Leu	Gln	His	Ala	Gly	Arg	Ile	Leu	Thr	Thr
				440					445					450
Arg	Ile	His	Glu	Leu	Lys	Gln	Ala	Lys	Val	Ser	Thr	Val	Met	Asp
				455					460					465
Thr	Val	Gly	Arg	Arg	Leu	Arg	Asp	Ser	Leu	Gly	Arg	Thr	Val	Asp
				470					475					480
Leu	Gly	Leu	Thr	Gly	Asn	His	Gly	Val	Val	His	Pro	Ser	Leu	Ala
				485					490					495
Asp	Ser	Ala	Asn	Lys	Phe	Glu	Glu	Asn	Thr	Tyr	Cys	Phe	Gly	Arg
				500					505					510
Thr	Val	Glu	Thr	Leu	Leu	Leu	Arg	Phe	Gly	Lys	Thr	Ile	Met	Glu
				515					520					525
Glu	Gln	Leu	Val	Leu	Lys	Arg	Val	Ala	Asn	Ile	Leu	Ile	Asn	Leu
				530					535					540
Tyr	Gly	Met	Thr	Ala	Val	Leu	Ser	Arg	Ala	Ser	Arg	Ser	Ile	Arg

Ile Gly Leu Arg	Asn His Asp His Glu	Val Leu Leu Ala Asn	Thr
545	550	555	
560	565	570	
Phe Cys Val Glu	Ala Tyr Leu Gln Asn	Leu Phe Ser Leu Ser	Gln
575	580	585	
Leu Asp Lys Tyr	Ala Pro Glu Asn Leu	Asp Glu Gln Ile Lys	Lys
590	595	600	
Val Ser Gln Gln	Ile Leu Glu Lys Arg	Ala Tyr Ile Cys Ala	His
605	610	615	
Pro Leu Asp Arg	Thr Cys		
620			

<210> 25
 <211> 245
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 5596281CD1

<400> 25

Met Gly Arg Leu Asp	Gly Lys Val Ile Ile	Leu Thr Ala Ala Ala	
1	5	10	15
Gln Gly Ile Gly Gln	Ala Ala Ala Leu	Phe Ala Arg Glu Gly	
20	25	30	
Ala Lys Val Ile Ala	Thr Asp Ile Asn	Glu Ser Lys Leu Gln	Glu
35	40	45	
Leu Glu Lys Tyr Pro	Gly Ile Gln Thr	Arg Val Leu Asp Val	Thr
50	55	60	
Lys Lys Lys Gln Ile	Asp Gln Phe Ala	Ser Glu Val Glu Arg	Leu
65	70	75	
Asp Val Leu Phe Asn	Val Ala Gly Phe	Val His His Gly Thr	Val
80	85	90	
Leu Asp Cys Glu Glu	Lys Asp Trp Asp	Phe Ser Met Asn Leu	Asn
95	100	105	
Val Arg Ser Met Tyr	Leu Met Ile Lys	Ala Phe Leu Pro Lys	Met
110	115	120	
Leu Ala Gln Lys Ser	Gly Asn Ile Ile	Asn Met Ser Ser Val	Ala
125	130	135	
Ser Ser Val Lys Gly	Val Val Asn Arg	Cys Val Tyr Ser Thr	Thr
140	145	150	
Lys Ala Ala Val Ile	Gly Leu Thr Lys	Ser Val Ala Ala Asp	Phe
155	160	165	
Ile Gln Gln Gly Ile	Arg Cys Asn Cys	Val Cys Pro Gly Thr	Val
170	175	180	
Asp Thr Pro Ser Leu	Gln Glu Arg Ile	Gln Ala Arg Gly Asn	Pro
185	190	195	
Glu Glu Ala Arg Asn	Asp Phe Leu Lys	Arg Gln Lys Thr Gly	Arg
200	205	210	
Phe Ala Thr Ala Glu	Glu Ile Ala Met	Leu Cys Val Tyr Leu	Ala
215	220	225	
Ser Asp Glu Ser Ala	Tyr Val Thr Gly	Asn Pro Val Ile Ile	Asp
230	235	240	
Gly Gly Trp Ser Leu			
245			

<210> 26
 <211> 159
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 5731013

<400> 26
 Met Thr Ala Arg Gly Thr Pro Ser Arg Phe Leu Ala Ser Val Leu
 1 5 10 15
 His Asn Gly Leu Gly Arg Tyr Val Gln Gln Leu Gln Arg Leu Ser
 20 25 30
 Phe Ser Val Ser Arg Asp Gly Ala Ser Ser Arg Gly Ala Arg Glu
 35 40 45
 Phe Val Glu Arg Glu Val Ile Asp Phe Ala Arg Arg Asn Pro Gly
 50 55 60
 Val Val Ile Tyr Val Asn Ser Arg Pro Cys Val Pro Arg Val
 65 70 75
 Val Ala Glu Tyr Leu Asn Gly Ala Val Arg Glu Glu Ser Ile His
 80 85 90
 Cys Lys Ser Val Glu Glu Ile Ser Thr Leu Val Gln Lys Leu Ala
 95 100 105
 Asp Gln Ser Gly Leu Asp Val Ile Arg Ile Arg Lys Pro Phe His
 110 115 120
 Thr Asp Asn Pro Ser Ile Gln Gly Gln Trp His Pro Phe Thr Asn
 125 130 135
 Lys Pro Thr Thr Phe Arg Gly Leu Arg Pro Arg Glu Val Gln Asp
 140 145 150
 Pro Ala Pro Ala Gln Val Gln Ala Gln
 155

<210> 27
 <211> 291
 <212> PRT
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 5731162CD1

<400> 27
 Met Ala Cys Leu Ser Pro Ser Gln Leu Gln Lys Phe Gln Gln Asp
 1 5 10 15
 Gly Phe Leu Val Leu Glu Gly Phe Leu Ser Ala Glu Glu Cys Val
 20 25 30
 Ala Met Gln Gln Arg Ile Gly Glu Ile Val Ala Glu Met Asp Val
 35 40 45
 Pro Leu His Cys Arg Thr Glu Phe Ser Thr Gln Glu Glu Glu Gln
 50 55 60
 Leu Arg Ala Gln Gly Ser Thr Asp Tyr Phe Leu Ser Ser Gly Asp
 65 70 75
 Lys Ile Arg Phe Phe Phe Glu Lys Gly Val Phe Asp Glu Lys Gly
 80 85 90
 Asn Phe Leu Val Pro Pro Glu Lys Ser Ile Asn Lys Ile Gly His
 95 100 105
 Ala Leu His Ala His Asp Pro Val Phe Lys Ser Ile Thr His Ser
 110 115 120
 Phe Lys Val Gln Thr Leu Ala Arg Ser Leu Gly Leu Gln Met Pro
 125 130 135
 Val Val Val Gln Ser Met Tyr Ile Phe Lys Gln Pro His Phe Gly
 140 145 150
 Gly Glu Val Ser Pro His Gln Asp Ala Ser Phe Leu Tyr Thr Glu
 155 160 165
 Pro Leu Gly Arg Val Leu Gly Val Trp Ile Ala Val Glu Asp Ala
 170 175 180
 Thr Leu Glu Asn Gly Cys Leu Trp Phe Ile Pro Gly Ser His Thr
 185 190 195

Ser	Gly	Val	Ser	Arg	Arg	Met	Val	Arg	Ala	Pro	Val	Gly	Ser	Ala
				200					205					210
Pro	Gly	Thr	Ser	Phe	Leu	Gly	Ser	Glu	Pro	Ala	Arg	Asp	Asn	Ser
				215					220					225
Leu	Phe	Val	Pro	Thr	Pro	Val	Gln	Arg	Gly	Ala	Leu	Val	Leu	Ile
				230					235					240
His	Gly	Glu	Val	Val	His	Lys	Ser	Lys	Gln	Asn	Leu	Ser	Asp	Arg
				245					250					255
Ser	Arg	Gln	Ala	Tyr	Thr	Phe	His	Leu	Met	Glu	Ala	Ser	Gly	Thr
				260					265					270
Thr	Trp	Ser	Pro	Glu	Asn	Trp	Leu	Gln	Pro	Thr	Ala	Glu	Leu	Pro
				275					280					285
Phe	Pro	Gln	Leu	Tyr	Thr									
				290										

<210> 28
 <211> 1557
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 543496CB1

<400> 28
 ctgagtgcga cggcgcaggt ctgcaccatg gcggcccggc ttgtcagccg atgcgggggt 60
 gtgcgtgcag ctccccacag cggcccgcgt gtgtcctggc gcaggtgggt cggcgcctca 120
 acagacaccg tgtatgacgt ggtgggtgtcg ggtggaggcc tgggtgggcgc tgccatggcc 180
 tgtgccttgg gatatgatata tcaactttcat gacaagaaaa tctgttgct cgaagcaggt 240
 ccaaagaaa tactggagaa attgtcagaa acttacagca acagggtcag ctccatttcc 300
 cctggctctg caacgcttct cagtagtttt ggtgcctggg accatatctg caacatgaga 360
 tacagagcct ttcggcgaat gcaggtgtgg gacgcctgct cagaggccct gataatgttt 420
 gataaggata atttagatga catgggctat atcgtggaga atgatgtcat catgcatgct 480
 ctactaagc agttggagge tgtgtctgac cgagtgcagg ttctctacag gagcaaaagg 540
 attcgtata cctggccttg tccatttct atggccgact ccagcccttg ggttcattt 600
 accctagggt atggcagcac ctccagacc aaattgttga taggtgcaga tggtcacaac 660
 tccggagtac ggcaggctgt tggaaatccag aatgtgagct ggaactatga ccagtctgct 720
 gttgtggcta ctctgcatt atcagaggcc acagaaaaca acgtagcctg gcagagattt 780
 ctccctctg ggcctattgc tctgtctccg ctctcagaca ccttgagttc cttggtttgg 840
 tccacgtccc atgaacatgc agcagagcta gttagcatgg atgaggaaaa atttgtggat 900
 gccgttaact ctgccttttg gagtgatgct gaccacacgg acttcacga cacagctggt 960
 gccatgctgc agtatgctgt cagccttctg aagcccacta aggtctcggc tcgccagctg 1020
 ccccaagcg tagccagggt ggatgccaaa agccgagttc tgtttcctct tgggttggga 1080
 catgctgctg agtacgtcag gcctcgggtg gcgctcattg gggatgcagc ccacagagtc 1140
 catccgcttg caggacaggg tgtcaacatg ggctttgggg atatctccag cttggcccat 1200
 cacctcagta cggcagcctt caatgggaag gacttaggtt ccgtgagcca cctcacaggt 1260
 tatgaaacag aaagacagcg tcacaacact gctcttctgg ctgctacaga ctactaaaa 1320
 aggctctatt ctaccagtgc ctccccgctt gtgttgctca ggacgtgggg cttgcaggcc 1380
 acaaatgcag tgtctccact caaagaacag attatggcct ttgcaagcaa atgagtactc 1440
 ctctcctaaa gaaagattac gttgatgaaa aagaacatcc tgcccaggac ccatcataca 1500
 tattttcaag atcttattta atttaataaaa cttactttac attaaaaaaa aaaaaaa 1557

<210> 29
 <211> 1106
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 907607CB1

<400> 29
 cccacgcgtc cgggttcccc ggctctcttt ggtcagggtg acgcagtagc ctgcaaacct 60

```

cgggcgctag gccaccgcac ttatccgcag caggaccgcc cgcagccggt aggggtgggct 120
cttcccagtg cccgcccagc taccggccag cctgcggtcg cgcagatctt tctgtggttct 180
gtcaggggaga cccttaggca ctccggacta agatggcggc gacggccagg cggggctggg 240
gagctgcggc tgttgccgcc gggtgcgca ggcggttctg tcatatgttg aagaatccat 300
acaccattaa gaaacagct ctgcatcagt ttgtacaaag accacttttc ccactacctg 360
cagcctttta tcaccagtg agatacatgt ttattcaaac acaagatacc ccaaatacaa 420
acagcttaaa gtttatacca ggaaaaccag ttcttgagac aaggaccatg gattttccca 480
ccccagctgc agcatttcgc tcccctctgg ctaggcagtt atttaggatt gaaggagtaa 540
aaagtgtctt ctttggaaca gatttcatca ctgtcacaaa ggaaaatgaa gaattagact 600
ggaatttact gaaaccagat atttatgcaa caatcatgga cttctttgca tctggcttac 660
ccctggttac tgaggaaaca ccttcaggag aagcaggatc tgaagaagat gatgaagtgt 720
tggaatgat taaggaattg ttagatacta gaatacggcc aactgtgcag gaagatggag 780
gggatgtaat ctacaaaggc tttgaagatg gcattgtaca gctgaaactc cagggttctt 840
gtaccagctg ccctagtcca atcattactc tgaaaaatgg aattcagaac atgtgcagt 900
tttatattcc ggaggtagaa ggcgtagaac aggttatgga tgatgaatca gatgaaaaag 960
aagcaaaactc accttaaaat aatctggatt ttctttgggc ataacagtca gacttggtga 1020
taatataat caagttttta ttattaatat gctgaggaac ttgaagatta ataaaaatat 1080
ctcttcagag aatgatatat aaaaaa 1106

```

<210> 30
 <211> 2180
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 1290078CB1

```

<400> 30
gccccacaca ggccgcggcg gctggctcgg gcccctacgg tcccggcggc ggctggagga 60
ggaagccagg cggctggcgg aggaggagag acggaggagg ccgagaccgg agcgccgctc 120
gccgcagact tacttccccg gctcagcagg gaaagggttc tagaagggtga gcgcggacgg 180
tatgcaaagt tgtgaatcca gtggtgacag tgccgatgac cctctcagtc gcgccctacg 240
gagaagggga cagcctcgtg tgggtggtgat cggcgccggc ttggctggcc tggctgcagc 300
caaagcactt cttgagcagg gtttcacgga tgtaactgtg cttgaggctt ccagccacat 360
cggaggccgt gtgcagagtg tgaacttgg acacgccacc tttgagctgg gagccacctg 420
gatccattgc tcccatggga accctatcta tcatctagca gaagccaacg gcctcctgga 480
agagacaacc gatggggaac gcagcgtggg ccgcacagc ctctattcca agaattggct 540
ggcctgctac cttaccaacc acggccgcag gatccccaag gacgtgggtg aggaattcag 600
cgattttatac aacgaggtct ataacttgac ccaggagttc ttccggcagc ataaaccagt 660
caatgctgaa agtcaaaata gcgtgggggt gttcaccgga gaggaggtgc gtaaccgcat 720
caggaatgac cctgacgacc cagaggctac caagcgctg aagctcgcca tgatccagca 780
gtacctgaag gtggagagct gtgagagcag ctcacacagc atggacgagg tgtccctgag 840
cgcttcggg gagtgaccg agatccccgg cgctcaccac atcatccct cgggcttcat 900
gcgggttgtg gagctgctgg cggagggcat ccctgccac gtcaccagc tagggaaacc 960
tgtccgtgc attcactgg accagcctc agcccgccc agaggccctg agattgagcc 1020
ccgggtgtag ggcgaccaca atcacgacac tggggagggg ggccaggggtg gagaggagcc 1080
ccggggggggc aggtgggatg aggatgagca gtggtcgggt gtggtggagt gcgaggactg 1140
tgagctgate ccggcggaac atgtgattgt gaccgtgtcg cttaggtgtg taaagaggca 1200
gtacaccagt ttcttccggc caggcctgcc cacagagaag gtggctgcca tccaccgct 1260
gggcattggc accaccgaca agatctttct ggaattcgag gagcccttct ggggcctga 1320
gtgcaacagc ctacagttt tgtgggagga cgaagcggag agccacaccc tcacctacc 1380
acctgagctc tggtagcgca agatctgcgg ctttgatgtc ctctaccgc ctgagcgcta 1440
cgcccatgtg ctgagcggct ggatctgcgg ggaggaggcc ctcgatcatg agaagtgtga 1500
tgacgagga gtggccgaga tctgcacgga gatgtgcgt cagttcacag ggaaccccaa 1560
cattccaaaa cctcggcgaa tcttgctcgc ggcctggggc agcaaccctt acttcgcgg 1620
ctcctattca tacacgcagg tgggtccag cggggcggat gtggagaagc tggccaagcc 1680
cctgccgtac acggagagct caaagacagc gcccatgcag gtgctgtttt ccggtgaggg 1740
caccacccgc aagtactatt ccaccacca cgggtgctct ctgtccggcc agcgtgaggg 1800
tgcccgctc attgagatgt accgagacct ctccagcag gggacctgag ggtgtctc 1860
gctgctgaga agagccacta actcgtgacc tccagcctg cccttgctgc cgtgtgctc 1920
tgcccttctg atcctctgta gaaaggattt ttatcttctg tagagctagc cgccctgact 1980
gccttcagac ctggccctgt agcttttctt tttctccagg ctggggcgggt agcaggtggg 2040
ccgttgagtt acctctgtgc tggatcccg gccccactt gcctaccctc tgtcctgct 2100
tgttattgta agtgccctca atactttgca ttttgggata ataaaaaagg ctccctcccc 2160
tgccctcaaa aaaaaaaaaa 2180

```

<210> 31
<211> 1311
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1302741CB1

<400> 31
ccgccgcccc cccgagccta tctgggctgc gtctttctgc cgctgctctt cgtggcccaa 60
cgccccaatc cttgcgtgtg cttgcagtc caccacacac tcagccttgt gtccctcgat 120
ccagtctccg acttccattt cccaccctaa accgcctacc cgggtgtctgt tccccgcccg 180
gttgctctcg ccctgctgcg ctgagtgtcc cctgttagcc tcgaccccat ggcgctgcag 240
acgctgcaga gctcgtgggt gaccttccgc aagatcctgt ctcaactccc cgaggagctg 300
agtctggctt tcgtctacgg ctccgggggtg taccgccagg cagggccgag ttcagaccag 360
aagaatgcta tgcgtggactt tgtgttcaca gtagatgacc ctgtcgcag gcattcaaag 420
aacctgaaga aaaattggag tcaactactt ttctataaag ttttagggcc caagattatc 480
acgtccatcc agaataacta tggcgctgga gtttactaca attcattgat catgtgtaat 540
ggtaggctta tcaaataatg agttatttag actaacgttc tgattgaaga tctcctcaac 600
tggaataact tatacattgc tggacgactc caaaaaccgg tgaaaattat ctcaagtgaac 660
gaggatgtca ctcttagatc agccctcgat agaaatctga agagtgtgt gaccgtgct 720
ttctctatgc tccccgaaag cttttctgaa gaagacctt tcatagagat tgccgggtctc 780
tctatttcag gtgactttcg gatgggtggt ggagaagata aaacaaaagt gttgaatatt 840
gtgaagcccc atatagccca ctttcgagag ctctatggca gcatactaca ggaaaatcct 900
caagtgggtg ataaaagcca gcaaggctgg ctggagatag ataaaagccc agaaggacag 960
ttcactcagc tgatgacatt gcccaaaacc ttacagcaac agataaatca tattatggac 1020
cctcctggaa aaaacagaga tgtggaagaa actttattcc aagtggctca tgatcccgac 1080
tgtggagatg tgggtgcgact agggctttca gcaatcgtga gaccgtctag tataagacag 1140
agcacgaaag gcatttttac tgcctggcctg aagaagtcag tgatttatag ttcactaaaa 1200
ctgcacaaaa tgtggaaagg gtggctgagg aaaacatcct gattttgctt gcttttatat 1260
atgttatatg tagatgaata aagtgtttga tcttttttga caaaaaaaaaa a 1311

<210> 32
<211> 921
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1541028CB1

<400> 32
ggccccctgc cccacgtgg cgccctttcc gggcgggggg ggtggggagg cacctccagc 60
gacaactctc cccctcccc cggcaccggg ctgatttgct gtgccactgg gaggggttcg 120
gggtggctga gctgagagg ctccgggaag gagtgaagtc aggacctgcc ttcagcccc 180
tttgtgccgc gggacgcccc gcggcacgac gtggagctcc tggagagcag ctccctgcc 240
atthttgccc acttcggggc tgtgagctgg gacttctcag ggactacccc tccgctcgg 300
ggctgggtcc cactctggc cccgggtgc taccagctcc tgctgtacca agcaggccgg 360
tgccaaccca gcaactgccg ccggtgcccg ggggcctatc gggcactgag ggggcttcga 420
agctttatga gtgccaacac cttcggaat gccggctttt ccgttctcct gcctggggcc 480
cggtctgagg gccgctgtgg gccaccaat gcccggttca gatgccatct gggcctaaag 540
atcctactct gctgtgagct ggtggtcggc ggtgagcccc agtgctgggc tgaggggac 600
tgtctactgg tggacgactc ttttctacac acagtggctc acaatggctc cccgaagat 660
gggcctcgag tggcttctcat cgtggacctc tggcaccaca acgtggcagg ggctgagcgc 720
caggccctcg actttgtctt cgccccagac ccttgaagga aggtgctccc ttcacacacc 780
caggctggag agacactgcy ctcagggacg gcttgatggt agccaggacc tctctctac 840
tgccgggggtg ggcggggggc gaggatggga actggctagt gagcactgaa atataaatc 900
tgaatcctca aaaaaaaaaa a 921

<210> 33
<211> 2032
<212> DNA
<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1597687CB1

<400> 33

```

agaaaccagg cggcgcggttc ccgggtggcgg cgccttgagc tcccggggccc ggcgatcccc 60
gccagccttc cttaaggcgg atgggtggcc cccgagaccc cgtcggaccc atgggttcca 120
gtgcagcgcg gagtgggcca tgccagcgtg ccaggagcca tgtctgacca ggacgtttgg 180
aagatcatat ccattgccaga ggctcttgtg aggagatgag ttggtaaaga gagaggctgg 240
gatgaagatg ttgccaggag tgggcgtggt tgggactggc agctccgccc gagttctggt 300
cccactgctg agggcagaag ggttactgtg tgaggccctg tgggggaaga ctgaggagga 360
ggcgaagcag cttgctgagg agatgaacat cgccttctac accagccgga ctgatgacat 420
cttgctgcat caagatgtgg atctgggtgtg catcagcatc cccctccac tcaccgggca 480
gatatccgtg aaggctctag gtattgggaa gaatgtggtt tgcgagaagg cagcaacatc 540
ggtggatgcc ttccggatgg tgacagctc cgcgtactac ccgcagctca tgagcctggt 600
agggaacgtg ctgcgcttcc tgctgcctt cgtgcgcatg aaacagctga tttcggaaca 660
ctatgtggga gcggtgatga tctgtgatgc ccgcatctac tcaggcagcc tgctgagccc 720
cagctatggc tggatctgtg atgagctcat gggcggcggg ggcttgaca ccatggggag 780
ctacattgtg gacctgtga cccacctgac cggccggaga gccgagaagg tgcacgggct 840
gctcaagaca ttctgtaggc agaacgtgc catccgtggc atccggcagc tctactagca 900
tgacttctgt ttcttccaga tgctcatggg tgggggtgtg tgtagcacag tgacactcaa 960
cttcaacatg ccaggcgctt ttgtgcatga agtcatggtg gtaggctctg caggacgctt 1020
cgtcggccgg ggagccgacc tctatgggca gaagaactct gccacgcaag aggagtgtct 1080
cttgaggtag ctgctggcag tgggcgcagg actgcctgag caggggcccc aggatgtccc 1140
gctgctgtac ctgaaggcca tgggtctacat ggtgcaggcc ttgcgcagc ccttccaggg 1200
gcagggcgac cgccgcacct gggaccgcac ccctgtctcc atggccgctt ccttcgagga 1260
tgggctgtac atgcagagcg tgggtggatgc catcaagagg tcgagccgat ccggggagtg 1320
ggaggctgtg gaggtgctga cggaggagcc cgacaccaac cagaacctgt gtgaggcact 1380
tcagcggaac aacctatgag cctgcacctg ggctccttgc cacagggcag agggaccagg 1440
gaggggaaca ggagccagac atgacaggga cttggcccta gcagagacag tgtctttcat 1500
ttaatgagtc tgggtgaagc cagagggtggc tctgggactt cgccccctcc agtggctcat 1560
ctgctcctca gacttgcagg gtggcaaatg ttctgttgc caggttacca cgggtgtcac 1620
aggccagagc tgccgccagg cctcctcatg gtgatgccag tgagtgggac caggaaaggc 1680
tcacctgggg gagcccttca gcctgatccg cagtccctaac aggggtgagaa agcacaagag 1740
gggcccagctc ggctgaggcc tgaggagaaa tgatgagaag ttttgtttcg ttcttgggct 1800
ggggcatctc aggaacagtg agggcaacaa gctgtccctc agggaccctc acctgcctc 1860
ccatgtgacc aggccttcc cagggtggga ctggggcttg ctttattaag tgagctcttg 1920
ggcttttgaa taggacaga gggccctggt ttggcagaga ggaggcaagg gagggcttta 1980
gaacatctga gaaatgttaa taaataatc agaaaataaa aaaaaaaaaa aa 2032

```

<210> 34

<211> 1134

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 1690348CB1

<400> 34

```

cgtggacttc taagaaagcg ccatggcctg tgctgctggt atgattcctg ggttgttgcg 60
gtgctctgtt ggagccatcc gtattgagcg tgcgtcactg agattgacac tcagcacttt 120
gcgccatctt actctaacca gcataatgaa atccaaaagg aaaactgatc acatggagag 180
aactgcaagt gtccttcgac gggagattgt ggcagcagct aagggtgtgtg gagctgccag 240
tgagtccagg tcagtgaaga gcctccgctt gcttgttctt gatcaagact ttccctttaa 300
agctggccag tgggttgatt tctttattcc aggtgtctct gtggttgggtg ggttttcaat 360
atgctccagt cccagactgc tagaacaaga gagagtgata gaattggcag tgaaatatac 420
gaaccaccct cctgccctct gggttcacaa tacgtgtaca cttgactgtg aagtggctgt 480
gagagtgggt ggagagttct tctttgaccc tcagcctgcg gatgcctcta gaaacctcgt 540
gttgattgca ggaggagtgc gaattaacct tctgctttcc atcctgcggc acgcagcaga 600
tctcctcaga gaggagcaa acaaaagaaa tggatatgag ataggacaa taaaactatt 660
ctacagtgca aaaaatacca gcgaactcct gtttaagaaa aatatccttg atttagtaaa 720
tgaatttctt gagaagattg catgcagttt gcatgttaca aaacagacta cacaatcaa 780
tgccggaactc aagccatata tcacggaagg aagaataacg gagaaggaga taagagatca 840
tatttcaaaa gagactttgt tctatatattg tggccacctt ccaatgacag acttttctc 900
caagcaactg gaaaacaacc atgtacccaa agaacacatt tgctttgaga agtgggtggt 960

```

```

ggaggcagac aaaggcagaa aaaataaaga ggtgagatct actcaggaga gctcctgtcc 1020
tttgtggcat gattaatttt ttttatctct acttgagttg tcttattttt taaggctata 1080
aacttagtga ccagctggat aataaaagcc agctggcaga cttaaatgat aaac 1134

```

```

<210> 35
<211> 734
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> unsure
<222> 661
<223> a or g or c or t, unknown, or other

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 1865603CB1

```

```

<400> 35
gggaggaccc tgggcaaaga cacctaccct gccatgctgc gccatctgcc ctccaggctg 60
ccagtcaaga tgtggggcag gactttggag aaacagtcac ggagggacag cagtcagacc 120
gtcccccat gtcttatccg tagacttgac cacatcgtga tgacggtgaa gagcatcaaa 180
gacaccacca tgtttttattc caagatcctg ggcattggagg tcatgacttt taaggaagac 240
cggaagacac tgtgttttgg agaccagaaa tttaacctcc acgaggtggg aaaggaattt 300
gaacccaaag ccgctcaccg agttcctggc tccctggaca tatgtctgat cacagaggtg 360
cctttggagg aaatgatcca gcacctcaag gcttgtgatg tccctattga ggaggggcca 420
gtccccagaa caggggcaaa agggcctatc atgtccatct acttccgaga ccccgacaga 480
aatctgattg aggtgtccaa ctacatctcc tcgtgatgga ggctggacct cctccattct 540
gtcccccttg atgtcgccct ctcccttcc tcttgcaaac cccacccag gccagagat 600
ctccaaagac tgaggactta ggcacttcac acatcctgct gagggggggac ccaagacagg 660
nttgggacca aacctacatg tctgtatgtc atcaaagttg gcctaataaa tgcataacct 720
ctcaatgaaa aaaa 734

```

```

<210> 36
<211> 2221
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<223> Incyte ID No: 1976472CB1

```

```

<400> 36
gtcggggggt tggcctctgc ccggccacag agccggagct ggaggtgctg tcccgctctgg 60
cggcgatccc cgggcagggc ccggggccgg ggtttgaaga tgctgaacgt cccttcccag 120
tctttcccgg cccccaggtc gcagcagcgt gtgcctccg gggggcgtag aaggtagctt 180
taaaacaggg cagaagcctt atggattgga ttcgactgac caaaagtgga aaggatctaa 240
cgggattaaa aggcagggtta attgaagtaa ctgaagaaga acttaagaaa cacaacaaaa 300
aagatgattg ttggatatgc ataagaggtt tctgttataa tctcagccct tatatggagt 360
atcatcctgg tggagaagat gaactaatga gagcagcagg atcagatggt actgaacttt 420
ttgatcaggt tcatcgttgg gtcaattatg aatccatgct gaaagaatgc ctggttggca 480
gaatggccat taaacctgct gttctgaaag actatcgtga ggagggaaaag aaagtcttaa 540
atggcatgct tcccagagc caagtgcagc atacacttgc caaagaaggt cctagttatc 600
caagctatga ttgggtccaa acagactctt tagtcaccat tgccatatat actaaacaga 660
aggatatcaa tttagactca ataatagttg atcatcagaa tgattccttt agagcagaaa 720
caattattaa ggattgttta tatcttatac atattgggct aagccatgag gttcaggaag 780
atttttctgt gcggtgtgtt gagagtgttg gaaaaataga gattgttcta caaaaaaaag 840
agaatacttc ttgggacttt cttggccatc cctgaagaa tcataattca cttattccaa 900
ggaaagatac aggtttgtac tacagaaagt gccaggtaat ttccaaggaa gatgttactc 960
atgatacgag gcttttctgt ttgatgctgc caccaagcac tcatcttcaa gtgcccattg 1020
ggcaacatgt ttacctcaag ctacctatta caggtagaca aatagtaaag ccatatacac 1080
ctgtactctg ttccttactc tcagagttca aggaaccagt tcttcccaac aataaataca 1140
tctacttttt gataaaaate tatcccactg gactcttcac accagagctt gatcgtcttc 1200
agattggaga ttttgtttct gtaagcagtc ctgagggcaa ttttaaaata tccaagttcc 1260
aagaattaga agatctcttt ttgttggcag ctggaacagg cttcacacca atggttaaaa 1320

```

```

tactgaatta tgctttgact gatataccca gtctcaggaa agtgaagctg atgttcttca 1380
ataaaaacaga agatgatata atttggagaa gccaatggga gaaattagca tttaaagata 1440
aaagactgga tggtgaattt gttctctcag cacctatttc tgaatggaat ggcaaacagg 1500
gacatatttc accagctctt ctttctgaat ttttgaaaag aaatttggac aaatccaaag 1560
ttctcgtctg catttgtgga ccagtgccat ttacagaaca aggagtaagg ttgctgcatg 1620
atctcaactt ttccaaaaat gagatccata gttttacagc ataatgaaga gctgtcattg 1680
tcctttatct aactagttta tctaaatttg tgattgctta ggggttttta agagaacatt 1740
tttgtagata acaaaagggt aactagaatc cagccttcag tttcttaaag gaaatcaaat 1800
gttccttcag tacaggtaac ttcttggtt tctttgtac cacaacttat ttactactg 1860
atatttgacc tggaaagtta atcatggcaa caaatacata caggattctt tgttatgaat 1920
cacaaatttc cttgccattt aaattatate actgttctac aataagcact tgtgttttat 1980
gacatgataa agtaaagtat cacttcgatc atgtttctat gctgtaaaat gtcttattag 2040
caatacagat taaattttac cttgcactgt taactcagga aatgatcatt tatgtccttg 2100
ttattaatgt taaaacatag aatgttataa cttattaagt gatccaaaca tttttttgtg 2160
tgtgtatggc attgatgcag aatagaataa aattatactt aagttctttt taaaaaaaaa 2220
a 2221

```

<210> 37

<211> 1706

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2050821CB1

<400> 37

```

agaagaggtt gtgtgggaca agctgtctcc gacagaagga tgtcgtctgt gagcctgccc 60
tggctgggcc tcagaccggt ggcaatgtcc ccatggctac tctgtctgtt ggttgtgggc 120
tcctggctac tcgcccgcct cctggcttgg acctatgcct tctataacaa ctgccgccgg 180
ctccagtgtt tcccacagcc cccaaaacgg aactggtttt ggggtcacct gggcctgate 240
actcctacag aggagggctt gaaggactcg acccagatgt cggccaccta ttcccagggc 300
tttacggtat ggctgggtcc catcatcccc ttcactgttt tatgccaccc tgacaccatc 360
cggctctatca ccaatgcctc agctgccatt gcacccaagg ataactctct catcagggtt 420
ctgaagccct ggctgggaga agggatactg ctgagtggcg gtgacaagtg gagccgccac 480
cgctggatgc tgacgcccgc cttccatttc aacatctga agtcctatat aacgatcttc 540
aacaagagtg caaacatcat gcttgacaag tggcagcacc tggcctcaga gggcagcagt 600
cgtctggaca tgtttgagca catcagcctc atgacctgg acagtctaca gaaatgcac 660
ttcagctttg acagccattg tcaggagagg cccagtgaat atattgccac catcttggag 720
ctcagtgcct ttgtagagaa aagaagccag catatctctc agcacatgga ctttctgtat 780
tacctctccc atgacgggcg gcgtctccac agggcctgcc gcctgggtgca tgacttcaca 840
gacgtgtca tccgggagcg gcgtcgcacc ctcccaactc aggggtattga tgatttttcc 900
aaagacaaag ccaagtccaa gactttggat ttcatgtatg tgcttctgct gagcaaggat 960
gaagatggga aggcatgttc agatgaggat ataagagcag aggctgacac cttcatgttt 1020
ggaggccatg acaccacggc cagtggcctc tcctgggtcc tgtacaacct tgcgaggcac 1080
ccagaatacc aggagcgctg ccgacaggag gtgcaagagc ttctgaagga ccgcatcct 1140
aaagagattg aatgggacga cctggcccag ctgcccttcc tgaccatgtg cgtgaaggag 1200
agcctgaggt tacatcccc agctcccttc atctcccgat gctgcaccca ggacattgtt 1260
ctcccagatg gccgagtcac ccccaaaggc attacctgcc tcatcgatat tatagggttc 1320
catcacaacc caactgtgtg gccggatcct gaggtctacg accccttcg ctttgaccca 1380
gagaacagca aggggaggtc acctctggct tttattcctt tctccgcagg gccccaggaa 1440
tgcatcgggc aggcgttcgc catggcggag atgaaagtgg tcctggcggt gatgctgctg 1500
cacttcgggt tcctgccaga ccacactgag ccccgagga agctggaatt gatcatgcgc 1560
gccgagggcg ggctttggct gcgggtggag cccctgaatg taggcttgca gtgactttct 1620
gacccatcca cctgtttttt tgcagattgt catgaataaa acggtgctgt caaaaaaaaaa 1680
aaaaaaaaac taaaaaaaaa aaaaaa 1706

```

<210> 38

<211> 549

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2408443CB1

<400> 38

```

cccacgcgtc cgtgtcatgt acttttgtta atatttttgc tataataaac ttagtatatt 60
tcaaggaggg caacatttta tggttacatt atactcgtct ccgagttgca cctcttgtcg 120
taaagcaaaa caatggctgg ttgaccataa tctcccattt attgaacgta atttaaataa 180
agaaccattg cgtgcggaag atgtcaaaag aatgttacga ttgactgaag atgggacgga 240
agaattaatt tcaacacgct caaaaatttt ttctgagttg acgattgact tagatgatat 300
gtcaattaat aaattgattg acctcatcgt catgtatccg tctttactga agcggccaat 360
tattcttgac gatcagcgca tgcaaatgg gtacaatgat gatgaaattc gtcgcttttt 420
accacgtgaa gttcgtcagc gagagttaat tcgcgcaaca tttaaagctg acttcgcaga 480
agaggcaaaa gatttagtgg ttgaagaagg ctgagcacgt cagtcttttt ttctttacca 540
tgacaaaag

```

<210> 39

<211> 1363

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2508668CB1

<400> 39

```

cgtcaggctc tgtgttgggt ggagcgagca tgtgggtctg cagtaccctg tggcgggtgc 60
gaacccccgc ccggcagtg cgggggctgc tccagcttc tggtgtcac ggacctgccg 120
cctcctccta ctccgcctcc gccgagcctg cccgggtccg ggcgttgtc tatgggcacc 180
acggggatcc agccaaggct gtcgaactca agaacctgga gctagctgct gtgagaggat 240
cagatgtccg tgtgaagatg ctggcggccc ctatcaatcc atctgacata aatatgatcc 300
aaggaaacta cggactcctt cctgaactgc ctgctgttgg agggaaacgaa ggtgttgac 360
aggtggtagc ggtgggcagc aatgtgaccg ggctgaagcc aggagactgg gtgattccag 420
caaatgctgg ttaggaacc tggcgggacc aggctgtgtt cagcgaggaa gcactgatcc 480
aagttccgag tgacatccct ctccagagcg ctgccacctt ggggtgcaat ccttgcacag 540
cctacaggat gttgatggac ttcgagcaac tgcagccagg ggattctgtc atccagaatg 600
catccaacag cggagtgggg caagcgggtc tccagatcgc cgcagccctg ggcctaagaa 660
ccatcaatgt ggtccgagac agacctgata tccagaagct gactgacaga ctgaagagtc 720
tgggggctga gcatgtcatc acagaagagg agctaagaag gcccgaaatg aaaaacttct 780
ttaaggacat gcccagcca cggttgtct tcaactgtgt tgggtgggaa agctccacag 840
agctgctgcg gcagttagcg cgtggaggaa ccatggtaac ctatgggggg atggccaagc 900
agcccgtcgt agcctctgtg agcctgtcct tttttaagga tctcaaactt cgaggctttt 960
ggttgtccca gtggaagaag gatcacagtc cagaccagtt caaggagctg atcctcacac 1020
tgtgcgatct catcccgca ggccagctca cagcccctgc ctgctcccag gtcccgtgc 1080
aggactacca gtctgccttg gaagcctcca tgaagccctt catatcttca aagcagattc 1140
tcaccatgtg atcatcccaa aagagctgga gtgacatggg aggggaggcg gatctgaggg 1200
gctgggtgca ggcccctcag ttggggctcc cacttcccc agactactgt tctcctcact 1260
gcctcttctt attaggagga tgggtgaagcc agccacgggt ttccccaggg ccagccttaa 1320
ggtatctaataaagtctgaa ctctcccttc caaaaaaaaaaaa aaa 1363

```

<210> 40

<211> 1196

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2536830CB1

<400> 40

```

ggttgccagg caaccgcgaa ccgcgggtccc tgcctctgag tctctctcga ccatggccaa 60
attcgtcatc gcgggtagag cagattgttc atattatgct aaaacagaaac ttgtggcaga 120
ctattttacaa agaattcttc ctgattttcg gatacataaa atcacacaac gtcctgaggt 180
ttgggaggat tggctaaaag atgtgtgtga aaagaataag tggagtcaca agaattcccc 240
tatcatctgg agagagctgt tggatcgtgg aggaaaggg ttgcttttgg gaggatataa 300
tgagtctctg gagcatgtc agctttacta tgatgtcacc tctagcatga cgactgaact 360
gatgatggta attgctcaag agaacctggg ggcacatata gaaaaagagc aggaggaaga 420
agccctgaaa acttgcata accccttgca ggtctggatc accagtgcct ctgctcctgc 480
ctgctacaac ctaattccca tattgacgag tggcgaagtgt tttgggatgc atacagaaat 540

```

tagcataact	ctatttgaca	acaagcaggc	ggaagaacat	ctcaaaagcc	ttgtggtgga	600
gacccaagac	ctggcatctc	ccgtcctgcg	cagtgtctcc	atctgcacga	aggtggagga	660
ggccttccgc	caggccacg	tcattgtggt	gctggatgac	agcaccaaca	aggaggtgtt	720
cactctggag	gactgcctcc	gaagcaggg	gcctctctgc	aggctctatg	ggtacctgat	780
agagaaaaat	gctcatgagt	ctgtcagagt	catcgtggga	gggagaacct	ttgtaaacct	840
gaagacagtt	ttactcatga	gatatgcccc	acgcattgca	cacaacatta	ttgctgtggc	900
gctgggggtg	gaaggtgaag	cgaaagccat	actggccaga	aaactgaaga	cagctccttc	960
atgtgagtga	atgtctcttt	atgtttttgt	ttttattttg	aaaacaacatg	atttcaaagc	1020
agctaggaag	aacactgtcc	cctctaggaa	cattctgagt	atctggaata	ggaagtataa	1080
aacagacata	aagaacttat	tgtgtgattt	gactactgtg	caaacatttt	tttttactgg	1140
cttttgtgtc	tggctttttc	catcattata	aaaatcaata	aatattttaat	aagaaa	1196

<210> 41

<211> 1926

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2645179CB1

<400> 41

gcttctcttc	ctggagtcca	gagtcctcgt	gctccgcgcg	gatattcagt	aaaccactgg	60
gagtcgagca	gcatggaggc	agcgcgccct	ccccgcagcg	caggggaagt	cgtggtgggc	120
ggcggcgcca	tcgcgggcgt	cacttgtgcg	gagcagttgg	ctactcactt	tccatcgga	180
gatattctct	tggtaacagc	ttctcctgtt	attaaagcag	ttacaaat	caagcagatt	240
tctaaaatat	tggagaatt	cgatgttgaa	gaacaatcaa	gtaccatgtt	aggaaaacgc	300
tttcccaaca	ttaaggttat	agaatctggc	gtaaagcaac	tgaagagtga	agaacactgc	360
atttgaacag	aagatggcaa	tcagcacgta	tataagaaac	tctgtctgtg	tgctggagct	420
aaaccaaagt	tgatatgtga	aggaaatcct	tatgtattag	gaatccgtga	tacagacagt	480
gctcaggaat	ttcagaaaca	gcttactaaa	gctaaaagaa	taatgatcat	agggaacggt	540
ggtattgcac	ttgagttagt	gtatgaaatt	gaaggctgtg	aagtgatttg	ggccattaaa	600
gataaagcta	tagggaatac	tttcttcgat	gcaggagcag	ctgaattcct	gacttcaaag	660
ctcattgctg	aaaaatcaga	ggctaaaatt	gcacataaaa	gaaccagata	tacaactgaa	720
ggaaggaaaa	aggaagctag	aagcaaactc	aaagcagata	atgtaggaag	tgcatgggga	780
ccagattggc	atgaaggctt	gaatcttaaa	ggaacaaaag	agttttctca	taagattcac	840
cttgaaacta	tgtgtgaagt	aaagaaaatc	taccttcagg	atgagtttag	aattttgaag	900
aaaaagtcct	tcacttttcc	aagagaccat	aagtcagtta	cagctgatac	agagatgtgg	960
cctgtctatg	tgggaattgac	caatgaaaag	atatatggct	gcgatttcat	tgctcagtgc	1020
acaggagtta	caccaaattgt	agaacctttt	ctccatggta	acagttttga	tctaggagaa	1080
gatggtggcc	tgaaagtggg	tgatcatatg	cacacatccc	ttcctgatat	ctatgctgcc	1140
ggtgacatct	tgactacatc	ctggcagctg	agcccagctc	ggcagcagat	gaggctgtgg	1200
accaggcta	gacagatggg	atggtatgca	gcaaagtgca	tggtctgcgc	gagttcagga	1260
gactctattg	acatggattt	cagctttgaa	ctgtttgctc	atgtgacaaa	attttttaac	1320
tataagggtg	tactgctggg	aaaatacaat	gcacagggct	taggttcaga	tcatgaatta	1380
atgctgagat	gtaccaaagg	acgagaatac	atcaaagtcg	tcatgcaaaa	tggacgaatg	1440
atgggagctg	tcttaattgg	tgaaaccgat	ttagaagaaa	catttgaaaa	cctaattctta	1500
aaccaaata	atctttcatc	atatggagaa	gatctgctag	atccaaatat	tgatatagaa	1560
gattattttg	actaaaaatg	gaatttcttc	aggaatcata	taaagttcca	aatgacacca	1620
gaaaaatcac	aagtcaataa	aatgaatgac	tgtattgagt	taatgatgac	cacactgaaa	1680
attacagaag	tgataatgat	attagtggaa	aaatataaaa	acataaattc	taagtttgaa	1740
atcagttcaa	agtttattta	tagatatatc	tttccaatac	aacactgacc	gcttagataa	1800
aaatcttaag	ttattttatt	ctgtgtttta	aacataaata	tgtttacttg	tgatttagct	1860
ttggagcaaa	tttaggttaag	ttatctactt	agccaaatgt	actctagtag	actagaacca	1920
ttcttt						1926

<210> 42

<211> 1727

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2754425CB1

<400> 42

```

agggaccgac ggcacttgat aatgtgacaa aaaccagcct cttccacccc actgcagtga 60
cttttgccgg aaagtggagc agtggggcag gccaaccttg tcgttcatca atgccgcccc 120
tggcagtatc tgaggatcgg tggcagccat gcctccccct gccccagccg ctctttcccc 180
cccacgctaa tctgcatggt gtggggcgccc ttggggacccc cgaggactga ttgtctgacc 240
ttgcttcaca cgcccagtaa ggactccccc aagatgtcgc tcgagtggct ggtggcctgg 300
agctggtcgc tggatggcct gagggactgc atcgccaccg gcatccagtc cgtgccccgg 360
tgcgacacca ccgctgtcat cactgtggcc tgccctcttg tccctcttctg gtggtagtgt 420
tatcacgtgg gcaggagaca gccccggccc tacgtctccg tcaactccct catgcaggct 480
gccgatgcca acgggctgca gaattggctac gtgtactgcc agtccccctga gtgcgtgcgc 540
tgcacccaca acgagggcct caaccagaag ctgtaccaca acctgcagga gtacgccaaag 600
cgctactcct ggtccggcat gggccgcata cacaagggca tccgcgagca gggccgggtac 660
ctcaacagcc ggccctccat ccagaagccc gaggtcttct tccctgcccga cctgcccacc 720
acgccttatt tctccccgga cgcacagaaa catgatgtgg aagtgtctga acggaaacttc 780
cagaccatcc tgtgtgagtt tgagaccctc tacaagctt tctcaaaactg cagcctcccc 840
caaggatgga aaatgaacag cccccccagc ggggagtggt tcacctttta cttggtcaat 900
caggggggttt gtgttcccag gaactgtagg aagtgtccac ggacgtaccg cttgctcgga 960
agccttcgga cctgtattgg gaacaatgtt ttggggaacg cgtgcatctc tgtgctgagc 1020
cctgggactg tgataacgga gcactatgga cccaccaaca tccgcatccg atgccattta 1080
ggtctgaaaa ctccaaatgg ctgtgagctg gtggtggggg gagagcccca gtgctgggca 1140
gaagggcgct gccttctctt tgatgactct ttcctgcatg ctgcgttcca tgaaggttca 1200
gcagaggatg gcccacgggt ggttttctat gtggatttgt ggcatccaaa cgtcgcagcg 1260
gccgaacggc aggtctctga ttccatcttt gctccgggac gatgagagta tttcccatgc 1320
tggagtccgc gagaaggggc gagggcgggc ctgggcagac tgtggtccgg tccagtccct 1380
accggtgttg tttccatgct cagaaacctg cctcagcgga aagctcttat ttgggatttt 1440
atatcatgtc gggctccctct ttccttgggt tattgtaaat ggaaactttt cggcttgtat 1500
ttccttagat tttttttttt tcttccaat catttgcttc agagactcct ttctggccta 1560
acagcgcat cctttgattg gtccttgagt gaccagagac ttagtgccct tgtaagtctg 1620
tcttctgttg ctacttgttt ttttcagtg cttgaaatag agtaactaaa tgggtatttg 1680
tctgaatata ataatgtaaa acttcttctg gtcactctta aaaaaaa 1727

```

<210> 43

<211> 611

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2821526CB1

<400> 43

```

cccacccggt ggggttttct cggcgccaccc gggaggcggg gccagcgagg caagatggag 60
ttagtgcagg tcttgaaacg cgggctgcag cagatcaccc gccacggcgg tctccgaggc 120
tatctacggg tttttttcag gacaaatgat gcgaagggtg gtacattagt gggggaagac 180
aaatatggaa acaaatacta tgaagacaac aagcaatttt ttggccgtca ccgatgggtt 240
gtatatacta ctgaaatgaa tggcaaaaac acattctggg atgtggatgg aagcatgggt 300
cctcctgaat ggcacgttg gcttcacagt atgactgatg atcctccaac aacaaaacca 360
cttactgctc gtaaatccat ttggacgaac cataaattca acgtgactgg caccacagaa 420
caatatgtac cttattctac cactagaaag aagattcagg agtggatccc accttcaaca 480
ccttacaagt aaagacaatg aagaacagtt gaaacatgca aaatatggag cttttcatgt 540
aattactctt ttactgttta ccattcacta taattcaca ttaaaattgt gtgactaaac 600
aaaaaaaaa a

```

<210> 44

<211> 1352

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 2876494CB1

<400> 44

```

gtttttattga caatacatgc atcatatctt ttgactttga aggatatctc atgtcaaagg 60
aatcaagtta tgatttatag aggattcagc tgggaatacct tgtgggtgct ggctgagggg 120
ggcaaaacgc ctaccgagac atgaagggtt tagccactag ttttgtcctt gggagcctgg 180
ggttggcctt ctacctgcct ttggtggtga ctacacctaa aacctggcc atccctgaga 240
agctgcaaga agctgtgggg aaagttatca tcaatgccac aacctgtact gtcacctgtg 300
gccttggcta taaggaggag accgtctgtg aggtgggccc tgatggagtg agaaggaaat 360
gtcagactca gcgcttagaa tgtctgacca actggatctg tgggatgctc catttcacca 420
ttctcattgg caaggaattt gagcttagct gtctgagttc agacatcttg gagtttggac 480
aggaagcttt ccggttcacc tggagacttg ctcgagggtg catctccact gacgatgagg 540
tcttcaaacc ctttcaagcc aactcccact ttgtgaagtt taaatatgct caggagtatg 600
actctgggac atatcgctgt gatgtgcagc tggtaaaaaa cttgagactt gtcaagaggc 660
tctatttttg gttgaggggc cttcctccta acttgggtgaa tctgaatttc catcagtcac 720
ttactgagga tcagaagtta atagatgagg gattggaagt taatctggac agctactcca 780
agcctcacca cccaaagtgg aaaaagaagg tggcgtcagc cttgggaata ggaattgcc 840
ttggagtggg ttggtggcgtg ttggtgagga ttgtcctctg tgcgctaagg gggggcctgc 900
agcagtgaca gcttcaagaa cttaacagcc ttgtcctga agaactggct gccaggaag 960
ccaagctagc tttttagggg agtgttccag ctgctggtag tggatcagct tagaggggaa 1020
actcccacag ccaaaagaat gagtgggaga aatggagggg acaatctcct gggagctatg 1080
cgcagtaacc taacttcctt atgtcccag gatctcttcc tgatcttccc tgcccattgg 1140
gtaccaggga aactgcaagc attgcctgtg ttcttgggaa gagttctaag aagcttgcag 1200
tcatttttcta ccttttatga cttggatgcc tccccacctc catttccctt cttctgagct 1260
gtgtattcat gtagagggat gtattcagcc tttttagtga acattttttt tcaataaaaag 1320
taattcacag tagctgtatg gagacctttt tc 1352

```

<210> 45

<211> 1458

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 3403225CB1

<400> 45

```

ggacggagag gagcagagaa gcagcagaag cagccaagag ctggagccag accaggaacc 60
tgagccagag ctgggggttga agctggagca gcagcaaaag caacagcagc tacagaagtt 120
ggaacgatgc tggtcacctt gggactgctc acctccttct tctcgttctt gtatatggta 180
gtcccatcca tcaggaagtt ctttgcctgt ggagtgtgta gaacaaatgt gcagcttctt 240
ggcaaggtag tggatgacac tggcgccaac acgggcatgt gcaaggagac ggccagagag 300
ctcgctagcc gaggagcccg agtctatatt gcctgcagag atgtactgaa gggggagttc 360
gctgccagtg aaatccgagt ggatacaaag aactcccagg tgctgggtgcg gaaattggac 420
ctatccgaca ccaaacttat ccgagccttt gctgagggct ttctggcaga ggaaaagcag 480
ctccatattc tgatcaacaa tgcgggagta atgatgtgtc catattccaa gacagctgat 540
ggctttgaaa cccacctggg agtcaaccac ctgggacacc gggtcaccac ctacgcagtg 600
caccaggcgg tcgtccgctc tgagctgggt cggcactcct cctgtctctg cctgtctctg 660
cggctcttct cccctttgt caagacggca cgggaggggg cgcagaccag cctgcactgc 720
gccctggctg agggcctgga gccctgagt ggcaagtact tcagtactg caagaggacc 780
tgggtgtctc caagggccc aaataacaaa acagctgagc gcctatggaa tgcagctgt 840
gagcttctag gaatccggtg ggagttagct gtggaagagc tgcagcttta tcaggcccaa 900
tccatgccat aatgaacagg gaccaaggag aaggccaacc ctaaaggatt gtcctcttgg 960
ccagctgggt ctgcgaatcc tgctgtctt ggcatgagtt acccttcttg gaattgttgc 1020
acacctgaca ctcttgtgag actggcttat ggcattgagtt gtggacacct atagagtgtt 1080
cttctctaag acctggaaag tcagcaacc tctgggggca gcaggactgg gcagatccca 1140
ggctgggcat ggggggtggca gaagagccc agaaaattggg tcagttccct catcagcacc 1200
agaggctcag ctgaggcaag aagagcacca tcaactgccta tttctagggg ctatacactc 1260
caactcttgg ttgatctctt tctttttaa aatatttgcc accacctgg agtctagacc 1320
aacacacaaa gatcctggct aaccttgccc tathtagatt ccttctctc accctggacct 1380
tcccatttca atcatgcaga tggtttctt ttgtaaagag ttccgtttgc ctttcaattt 1440
ttagagaaaa taaagact 1458

```

<210> 46

<211> 1884

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 4163943CB1

<400> 46

```

ggactccaag cgccatggcc gctgcccggc gagcccgggt cgcgtacttg ctgaggcaac 60
tgcaacgcgc agcgtgccag tgcccaactc attctcatac ttactcccaa gccctgggac 120
tttcaccttc tgggaaaaca acagattatg cctttgagat ggctgtttca aatattagat 180
atggagcagc agttacaaag gaagtaggaa tggacctaaa aaacatgggt gctaaaaatg 240
tgtgcttgat gacagacaag aacctctcca agctccctcc tgtgcaagta gctatggatt 300
ccctagttaa gaatggcatc ccctttacgg tttatgataa tgtgagagtg gaaccaacgg 360
attcaagctt catggaagct attgagtttg cccaaaaggg agctttttgat gcctatgttg 420
ctgtcgggtg ttgctctacc atggacacct gtaaggctgc taatctgtat gcatccagcc 480
ctcattctga ttctctagat tatgtcagtg ccccatggg caagggaaaag cctgtgtctg 540
tgctctttaa gctcttgatt gcagtgccaa ctacctcagg aaccgggagt gaaactactg 600
gggttgccat ttttgactat gaacacttga aagtaaaaaat tggcatcact tcgagagcca 660
tcaaaccac actgggactg attgatectc tgcacaccct ccacatgcct gcccgagtgg 720
tcgccaacag gtgctttgat atgccttggg atgccttggg gtcatacacc accctgccct 780
accacctgcg gagccctgc ccttcaaatc ccacacacg gctgcgtac cagggcagca 840
acccaatcag tgacatttgg gctatccacg cgctgcggat cgtggctaag tatctgaaga 900
gggctgtcag aaatcccgat gatcttgaag caaggtctca tatgcaactg gcaagtgcct 960
ttgctggcat cggctttgga aatgctggtg ttcactctgt ccattggaatg tcttacccaa 1020
tttcaggttt agtgaagatg tataaagcaa aggtattaca tgtggatcac ccactgggtc 1080
cccatggcct ttctgtggtg ctacagtcct cagcgggtgt cactttcag gcccgatgt 1140
ttccagagcg acacctggag atggcagaaa tactgggagc cgacaccgc actgccagga 1200
tccaagatgc agggctgggtg ttggcagaca cgctccggaa attcttattc gatctggatg 1260
ttgatgatgg cctagcagct gttggttact ccaaagctga tatccccgca ctagtgaag 1320
gaacgctgcc ccaggaaaagg gtcaccaagc ttgcaccccg tccccagtc gaagaggatc 1380
tggctgctct gtttgaagct tcaatgaaac tgtattaatt gtcattttta ctgaaagaat 1440
taccgctggc cattgtagt ctgagagcaa gagctgatct agctagggtt ttgtcttttc 1500
atctttgtgc ataacttacc tgttaccagt ataggtggga tatacattta tcttgcagga 1560
aattccccaa agctcagagt ccagttcctt ccataaaaaca ggctggacaa atgaccacta 1620
tgtagacccc ccaggctcga cttcaggggt cagtgttcct gtcaccaacc ccacacagaa 1680
tactctgcct ctgcttcag tagcaaatga gcaaaaactc agtatctatc aaaagtgtaa 1740
attatatttc ctatgcctag taattcactt catgtctaaa aatttatctg atagaaacac 1800
tagcaccagt acatacagaa gcattggcaag gatgtttctg gcagcacttt tctaataata 1860
aaagatttga aacaaaaaaa aaaa

```

<210> 47

<211> 1400

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 4293484CB1

<400> 47

```

ctccaggcca gtctgggac cctgggatag cggctgcagc tcccaggatc agctctgccc 60
tcccgccaaa cgccagcctc gtcaccgctc cagggcacct ccagcagtaa caggtgggtg 120
cagcaggtgg cagccagccc ctggatgagc caaggtctct tcccagcca ggcattggccg 180
actctgcaca ggccagaag ctggtgtacc tggtcacagg ggcgtgtggc ttcctggggag 240
agcagctggt gcgaatgctg ctgcagcggg agccccggct cggggagctg cgggtctttg 300
accaacacct gggctccctg ctggaggagc tgaagacagg taccgggaac gtgatcgagg 360
cttgtgtgca gaccggaaca cggttcctgg tctacaccag cagcatggaa gttgtggggc 420
ctaacaccaa aggtcacccc ttctacaggg gcaacgaaga caccatac gaagcagtg 480
acaggcaccc ctatccttgc agcaaggccc tggccgagtg gctggtcctg gaggccaacg 540
ggaggaaggc ccgtgggggg ctgcccttgg tgacgtgtgc ccttcgtccc acgggcatct 600
acgggtgaagg ccaccagatc atgaggact tctaccgcca gggcctgcgc ctgggaggtt 660
ggctcttcog ggccatcccg gcctctgtgg agcatggccg ggtctatgtg ggcaatgttg 720
cctggatgca cgtgctggca gccggggagc tggagcagcg ggcaaccctg atgggcggcc 780
aggtatactt ctgctacgat ggatcacctc acaggagcta cgaggatttc aacatggagt 840
tctggggccc cgtcggactg cggtctgttg ggcggcgccc attgtgccc tactggctct 900
tgggtgttcc ggctgcccct aatgccctgc tgcagtggct gctgcggcca ctgggtctct 960
acgcacccct gctgaacccc tacacgttgg ccgtggccaa caccaccttc accgtcagca 1020
ccgacaaggc tcagcgccat ttcggctatg agccccgtgt ctcgtgggag gatagccgga 1080

```

```

ccccgacccat tctctgggta caggecgcta cgggttcagc ccagtgacgg tggggctggg 1140
gcctggaggc ccagatacag cacatccacc cagggtccga gccctcacac cctggacggg 1200
aagggacagc tgcattccag agcaggaggc agggctctgg ggccagaatg gctgtccttg 1260
tcgtagagcc ctccacattt tctttttctt ttttgagaca gggctcttgct ctgtcaccca 1320
gactggagtg cagtgggtg atcatagctc actgcaccct caacctcctg ggttcaagca 1380
atcctcctgc ctccagctcc

```

<210> 48
 <211> 1313
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 4440080CB1

```

<400> 48
gaggctggct agggagccgt gggagtgage tgctcacccc agtcagaccc ctgttggggtt 60
tcataggagc tggctgctga atgtaagagt gcaggctacc ccgggacttt gatccccctac 120
agatgtgacc tatcaaatga agaggacatc ctctccatgt tctcagctat ccgttctcag 180
cacagcgggtg tagacatctg catcaacaat gctggccttg cccggcctga caccctgctc 240
tcaggcagca ccagtgggtg gaaggacatg ttcaatgtga acgtgctggc cctcagcatc 300
tgcacacggg aagcctacca gtccatgaag gagcggaatg tggacgatgg gcacatcatt 360
aacatcaata gcatgtctgg ccaccgagtg ttaccctgt ctgtgaccca cttctatagt 420
gccaccaagt atgccgtcac tgcgctgaca gagggactga ggcaagagct tggggaggcc 480
cagacccaca tccgagccac gtgcatctct ccagggtgtg tggagacaca attcgccttc 540
aaactccacg acaaggaccc tgagaaggca gctgccacct atgagcaaat gaagtgtctc 600
aaacccgagg atgtggccga ggctgttate tacgtcctca gcacccccgc acacatccag 660
attggagaca tccagatgag gccacggag cagggtgacct agtgactgtg ggagctcctc 720
cttccctccc cacccttcat ggcttgctc ctgcctctgg attttagggt ttgatttctg 780
gatcacggga taccacttcc tgtccacacc ccgaccaggg gctagaaaat ttgtttgaga 840
tttttatatc atcttgtcaa attgcttcag ttgtaaatgt gaaaaatggg ctggggaaag 900
gagggtgggt ccctaattgt tttacttgtt aacttgttct tgtgccctg ggcacttggc 960
ctttgtctgc tctcagtgtc ttccctttga catgggaaag gagttgtggc caaaatcccc 1020
atcttcttgc acctcaacgt ctgtggctca gggctggggg ggcagaggga ggccttcacc 1080
ttatatctgt gttgttatcc agggctccag acctcctcct ctgcctgccc cactgcacc 1140
tctccccctt atctatctcc ttctcggtc cccagcccag tcttggcttc ttgtccctc 1200
ctgggggtcat cctccactc tgactctgac tatggcagca gaacaccagg gcttggccca 1260
gtggatttca tggatgatcat taaaaaagaa aaatcgcaac caaaaaaaaaa aaa 1313

```

<210> 49
 <211> 1459
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <223> Incyte ID No: 5495687CB1

```

<400> 49
gggcccgcgc cgcgtcccca ggctctcttc ccccagcctt cctccggctg gcagcacgac 60
tcgcgtagcc gtgcgccgat tgcctctcgg cctgggcaat ggtcccggt gccggtcgac 120
gaccgccccg cgtcatgcgg ctccctcggc ggtggcaagt attgctgtgg gtgctgggac 180
ttcccgtccg cggcgtggag gttgcagagg aaagtggctg cttatgggtcaggaggcagc 240
ctgtctaccc tctccagggt ggggctgtgt acctgggtga ggaggagctc ctgcatgacc 300
cgatgggcca ggacagggca gcagaagagg ccaatgcggt gctggggctg gacacccaag 360
gcgatcacat ggtgatgctg tctgtgattc ctggggaagc tgaggacaaa gtgagttcag 420
agcctagcgg cgtcacctgt ggtgctggag gagcggagga ctcaagggtg aacgtccgag 480
agagcctttt ctctctggat ggcgctggag cacacttccc tgacagagaa gaggagtatt 540
acacagagcc agaagtggcg gaactctgac cagccccgac agaggactcc aataacactg 600
aaagtctgaa atccccaaag gtgaactgtg aggagagaaa cattacagga ttagaaaatt 660
tcactctgaa aattttaaat atgtcacagg accttatgga ttttctgaac ccaaacggta 720
gtgactgtac tctagtctct ttttacaccc cgtgggtgccc cttttctgcc agtttggccc 780
ctcactttta ctctctgccc cgggcatttc cagctcttca ctttttggca ctggatgcat 840
ctcagcacag cagcctttct accaggtttg gcaccgtagc tgttctaat attttattat 900

```

ttcaaggagc	taaaccaatg	gccagattta	atcatacaga	tcgaacactg	gaaacactga	960
aaatcttcat	ttttaatcag	acaggatatg	aagccaagaa	gaatgtgggtg	gtaactcaag	1020
ccgaccaaat	aggccctctt	cccagcactt	tgataaaaag	tgtggactgg	ttgcttgtat	1080
tttcttattt	ctttttaatt	agttttatta	tgtatgctac	cattcgaact	gagagtattc	1140
ggtggctaatt	tccaggacaa	gagcaggaac	atgtggagta	gtgatgggtct	gaaagaagtt	1200
ggaaagagga	acttcaatcc	ttcgtttcag	aaattagtgc	tacagtttca	tacattttct	1260
ccagtgcagt	gttgacttga	aacttcaggc	agattaaaag	aatcatttgt	tgaacaactg	1320
aatgtataaa	aaaattataa	actgggtgtt	taactagtat	tgcaataagc	aaatgcaaaa	1380
atattcaata	gatgcactat	tcttggtttt	actgcatgaa	cgtaattcag	tatttgga	1440
gtaatccagt	ttgaaatgg					1459

<210> 50

<211> 2101

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 5527735CB1

<400> 50

cggtctgagg	cggccgtcat	ggcgtcgccc	ttcagcgggg	cgctgcagct	gacggacctg	60
gatgacttca	tcgggccgtc	tcaggagtgc	atcaagcctg	tcaaagtgga	aaaaagggcg	120
ggaagtggcg	tggccaagat	tcgcattgaa	gatgacggga	gctacttcca	aattaacca	180
gacggcgagg	cccggaggct	ggagaaggcc	aaggtctcgc	taaacgactg	cctggcgtgc	240
agcggctgca	tcacctccgc	agagaccgtg	cttatcaccc	agcagagcca	cgaggagctg	300
aagaaggttc	tagatgctaa	caagatggcg	gcacccagtc	agcagaggct	ggttgtagtt	360
tcggtctcac	cacagtctag	agcatcgctg	gctgcacggg	ttcagctgaa	tcctacagat	420
actgccagga	aattaacctc	attctttaaa	aaaatagggg	tgcacttcgt	cttcgcacacc	480
gcctttctca	ggcacttcag	cctcctggag	agccagcgag	agtttgtgcg	gcgattccga	540
ggacaggccg	actgcagaca	ggcactgccc	ctgctggcct	ctgcctgccc	aggctggatc	600
tgctatgccg	agaagactca	cggcagcttc	atcctccccc	acatcagcac	cgcccgggtc	660
ccgcagcagg	tcattgggtc	cctggtcaag	gacttcttcg	cccagcagca	gcacttgacc	720
cctgacaaga	tctaccacgt	cacagtgatg	ccctgctatg	acaaaaagct	ggaagcctcc	780
agacccgact	ttttcaacca	ggagcaccag	acacgggatg	tggactgtgt	cctcacaaca	840
ggagaagttt	tcaggttgct	ggaggaagag	ggcgtctccc	tccccgacct	ggaaccagcc	900
ccctctggaca	gcctgtgcag	cgggtgcctc	gcagaggagc	ccaccagcca	tcggggaggg	960
ggctcggggg	gctacctgga	gcattgtgtc	cggcacgcgg	cccagagact	ctttggaatc	1020
catgtggctg	aggttacctt	caaaccctcg	aggaacaaaag	acttccagga	ggtgacactg	1080
gagaaggagg	gccagggtgc	gctgcacttc	gcaatggcgt	acggcttccg	caacatccag	1140
aacctgggtg	agaggctcaa	acgaggggcg	tgcccctacc	actacgtgga	ggtcatggcc	1200
tgcccctcag	gctgcctgaa	cggcgggggg	cagctccagg	ccccagacag	gcccagcaga	1260
gagctcctcc	agcacgtgga	gagactgtac	ggcatggtcc	gggctgaggc	gcccagaggac	1320
gcgcctgggg	ttcaggagct	gtacacacac	tggtgcagg	gcacggactc	ggagtgtgca	1380
ggctcgcttg	tgcatacgca	gtaccacgcc	gtggagaagg	ccagcactgg	cctgggcatc	1440
cgggtggtagg	cggtgcagga	ccaggactcc	caggaggccg	tgtccatgtg	tgacagcaga	1500
accacatgcc	ccaagacccc	agggcttccc	ccaaaattct	gagtgaagctg	cagggtgtgc	1560
tgggaccoga	gtaggagcta	ggactagcca	ggaccgcag	ccgcctcgtc	acctccagtt	1620
gggtgcctct	gggttcccac	tggtctctgc	caggtggggg	gggggtggcc	aggcagcaga	1680
aggttccctg	aggtcccaga	gcctgttccg	ttggccctgg	gccgaggccc	acaggtgctg	1740
cccttgctgc	tgctgggtcg	gcacccaagt	gcgtgagggg	cttcagcctg	tcccgggggt	1800
gcctgaggca	gagcaagacg	ggttctcacc	cctgacttct	ggaggcttcc	cttgaagctc	1860
tgtgcaaaag	gtgggagaca	gagctggacc	tgcaggggtg	gtcccggcac	aaccctgcgt	1920
gtggaccctg	gcaggggggg	ggtgccaggc	ccctggaaaag	caggggttac	cgttacgagg	1980
ctgtgggtccg	gggcaagcca	agtagcaagc	agcagccatc	gcgggctgca	tcattccccca	2040
gccagggtccc	caccaggcct	gtctcccagc	gtttgtctaa	taaacgcacc	cctcgtaaaa	2100
g						2101

<210> 51

<211> 2440

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 5540437CB1

<400> 51

gtgtccctgc	ggcgctaaga	aggggagact	gaggctgagg	ctgggggaaca	tcgggcagca	60
tgagcggctg	cgggctcttc	ctgcgcacca	cggtgcggc	tcgtgcctgc	cggggtctgg	120
tggtctctac	cgcgaaacgg	cggtacttgc	gcaccagccc	gcctgtacga	gctttcgcca	180
aagagctttt	cctaggcaaa	atcaagaaga	aagaagtttt	cccatttcca	gaagttagcc	240
aagatgaact	taatgaaatc	aatcagttct	tgggacccgt	ggaaaaatc	ttcactgaag	300
aggtggactc	ccgaaaaatt	gaccaggaag	ggaaaaatccc	agatgaaact	ttggagaaat	360
tgaagagcct	agggcttttt	gggctgcaag	tcccagaaga	atatgggtggc	ctgggcttct	420
ccaacacccat	gtactcaaga	ctaggggaga	tcatacagcat	ggatgggtcc	atcactgtga	480
ccctggcagc	gcaccaggct	attggcctca	aggggatcat	cttggctggc	actgaggagc	540
agaaagccaa	atacttgccct	aaactggcgt	ccggggagca	cattgcagcc	ttctgcctca	600
cggagccagc	cagtggggagc	gatgcagcct	caatccggag	cagagccaca	ctaagtgaag	660
acaagaagca	ctacatcctc	aatggctcca	aggtctggat	tactaatgga	ggactggcca	720
atattttttac	tgtgttttga	aagactgagg	tcgttgattc	tgatggatca	gtgaaagaca	780
aaatcacagc	attcatagta	gaaagagact	ttgggtggag	cactaatggg	aaacccgaag	840
ataaattagg	cattcggggc	tccaacactt	gtgaagtcca	ttttgaaaac	accaagatac	900
ctgtggaaaa	catccttgga	gaggtcggag	atgggttta	ggtggccatg	aacatcctca	960
acagcggccg	gttcagcatg	ggcagcgtcg	tggctgggct	gctcaagaga	ttgattgaaa	1020
tgactgctga	gtacgcctgc	acaaggaaac	agttaaaca	gaggctcagt	gaatttggat	1080
tgattcagga	gaaatttgca	ctgatggctc	agaaggctta	cgatcaggag	agtatgacct	1140
acctcacagc	agggatgctg	gaccaacctg	gctttcccg	ctgctccatc	gaggcagcca	1200
tggtgaaggt	gttcagctcc	gaggccgcct	ggcagtggtg	gagtgaggcg	ctgcagatcc	1260
tcgggggctt	gggctacaca	agggactatc	cgatcagcgc	catactgcgt	gacaccgcca	1320
tcctctcat	cttcgaggga	accaatgaga	ttctccggat	gtacatcgcc	ctgacgggtc	1380
tgcagcatgc	cggccgcate	ctgactacca	ggatccatga	gcttaaacag	gccaaagtga	1440
gcacagtcac	ggataccgtt	ggccggaggc	ttcgggactc	cctgggccga	actgtggacc	1500
tggggctgac	aggcaaccat	ggagtgtgtc	accctagtct	tgccggacagt	gccaaacagt	1560
ttgaggagaa	cacctactgc	ttcggccgga	ccgtggagac	actgctgctc	cgctttggca	1620
agaccatcat	ggaggagcag	ctggtactga	agcgggtggc	caacatcctc	atcaacctgt	1680
atggcatgac	ggcctgtctg	tcgcccggca	gccgtccat	ccgcattggg	ctccgcaacc	1740
acgaccacga	ggttctcttg	gccaacacct	tctgctggga	agcttacttg	cagaatctct	1800
tcagctctc	tcagctggac	aagtatgtct	cagaaaacct	agatgagcag	attaagaaag	1860
tgtcccagca	gateccttga	aagcgagcct	atatctgtgc	ccaccctctg	gacaggacat	1920
gctgaggcag	gggacagtgt	cccctgtctc	cgcccgcctc	taccatggc	ccgttgctgg	1980
atgactgtta	ctcttttttc	agaaggtgtt	gggattatca	caggttaagc	cttttgttcc	2040
ccgtctgcac	ctgaagggtt	gtgcctggc	ctgggagagc	ctcttccagg	ttttgacctg	2100
caggcagtcg	tctctaacag	gacctacaca	gcttctgaac	tgagccggag	agagagaatg	2160
gaattgctga	cccctggaac	tggcgggtat	tctggtcatt	gaggagacac	catagtggaa	2220
actggggctt	atgctgctgc	ctccagggtg	tgaggtgggt	ggggacctgt	gtcaggtgtg	2280
gatagccatt	tctgtcaaac	cacacattct	ctaagaaaca	gcttgaaagc	tctgtctggg	2340
tcattcattt	aaactagaag	cagaggcact	taaaacatgt	accaggaacc	atttaacaaa	2400
gaatataaaa	tgtcacaaac	tgtgtactgt	taaaaaaaaa			2440

<210> 52

<211> 1072

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 5596281CB1

<400> 52

cggacgggtg	gtccagacaa	aggttgtctc	aagtttgttg	ctcaaaccga	gttctggaga	60
acgccatcag	ctcgtgctt	aaaattaaac	cacaggttcc	attatgggtc	gacttgatgg	120
gaaagtcac	atcctgacgg	ccgtgtctca	ggggattggc	caagcagctg	ccttagcttt	180
tgcaagagaa	ggtgccaaag	tcatagccac	agacattaat	gagtccaaac	ttcaggaact	240
ggaaaagtac	ccgggtattc	aaactcgtgt	ccttgatgtc	acaaagaaga	aacaaattga	300
tcagtttgcc	agtgaaagttg	agagacttga	tgttctctt	aatgttgctg	gttttgcctg	360
tcattggaact	gtcctggatt	gtgaggagaa	agactgggac	ttctcgatga	atctcaatgt	420
gcgacagcatg	tacctgatga	tcaaggcatt	ccttccataa	atgcttgctc	agaaatctgg	480
caatattatc	aacatgtctt	ctgtggcttc	cagcgtcaaa	ggagttgtga	acagatgtgt	540
gtacagcaca	accaaggcag	ccgtgattgg	cctcacaana	tctgtggctg	cagatttcat	600
ccagcagggc	atcaggtgca	actgtgtgtg	cccaggaaca	gttgatacgc	catctctaca	660

agaaagaata	caagccagag	gaaatcctga	agaggcacgg	aatgatttcc	tgaagagaca	720
aaagacggga	agattcgcaa	ctgcagaaga	aatagccatg	ctctgcgtgt	atttggcttc	780
tgatgaatct	gcttatgtaa	ctggttaacc	tgatcatcatt	gatggaggct	ggagcttggt	840
attttaggat	ctccatgggt	ggaaggaagg	caggcccttc	ctatccacag	tgaacctggt	900
tacgaagaaa	actcaccaat	catctccttc	ctgttaataca	catgttaatg	aaaataagct	960
ctttttaatg	atgtcactgt	ttgcaagagt	ctgattcttt	aagtatatta	atctctttgt	1020
aatctcttct	gaaatcattg	taaagaaata	aaaatattga	actcaaaaaa	aa	1072

<210> 53

<211> 1040

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 5731013CB1

<400> 53

ccttcgggctc	gagcggcaac	cggaagcctc	gaggtttagt	cccgccccc	tctcctcgct	60
gcttaggctc	cgcgccctcc	aagctgtagc	tatgacggcg	cgcgggactc	cgagccgctt	120
cttggccagc	gttctccaca	acggactggg	tcgctatgtg	cagcagctgc	agcgtctgag	180
cttcagcgct	agcccgagcg	gcgcctcgtc	tcgcggcgcc	agggagttcg	tggagcggga	240
ggtagtcgac	ttcgcccagc	ggaatccagg	ggtcgtaata	tatgtaaact	cgcgctccgtg	300
ctgcgtgccc	agagtagtgg	ccgaatacct	taacggggct	gtgcgcgagg	agagcatcca	360
ctgcaagtgc	gtcgaggaga	tctcgacgct	ggtgcagaag	ctggccgacc	agtcgggctt	420
ggacgtgatc	cgcattccga	agcccttcca	caccgacaac	cctagcatcc	agggccagtg	480
gcaccccttc	accaacaagc	cgaccacggt	ccgcgggcta	cgccccgag	aggttcagga	540
tcctgcccc	gcccaggtgc	aagcacagtg	aagagttgcc	ccaccaactg	cagccccagg	600
ctttggactg	ttactccggt	aaagggtggt	cttccccctt	gggattccaa	gcccaggcaa	660
atggaaccca	tcaatgggca	agttgacaga	ggttctgctt	gggataatga	agagctgcct	720
gtttcttttc	agtgcctgct	tctgggggca	gtgacctgtg	gaaccactca	tttttatcga	780
agtggcatcc	ctaaaacctg	agatgaggaa	gacttcaagg	gttttacagg	acccttggtt	840
tttaaatacca	aattgataat	aatgatctca	aaacacagtg	agaggtctga	aggctggctt	900
ctgaagaatc	cctgatgtct	tattggaaca	accactgagc	tacggagagc	tctgctgtga	960
tgggctaggc	actttatatc	tgtgtgaata	cagatttata	aaacagggtta	ataaacttat	1020
ccaagggtcaa	aaaaaaaaaa					1040

<210> 54

<211> 1624

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<223> Incyte ID No: 5731162CB1

<400> 54

ctcagcctcg	cgagcagctg	ggactacagg	tgcccgccac	cacgcccggc	taattttttg	60
tatttttagt	agagacgggg	tttcaccgtg	ttagccagga	tggctctgat	ttcctgacct	120
cgtgatccgc	gcgcctctgc	ctccaaagtg	ctgggattac	aggcccgagc	cgccgcaccc	180
ggcctgcagg	tgttggtctg	acccagctc	cactgtgcca	gcccgaactg	agatgccaag	240
tatcttgggc	ctgaggattg	aggcctagag	agaggcaggc	gttcctcaga	gtcacaggga	300
atggcggcgc	ctggactggg	actcagccca	gctgcttggc	ctgacctctc	cacagcataa	360
tttcccgcca	cctggaggcc	gcgcttagaa	gccgcccagt	gccctgagcg	tctccatggc	420
ctgcctgagc	ccctcgagc	tccagaagtt	ccaacaggat	ggattcctgg	tgctggaagg	480
attcttgtct	gcggaagagt	gtgtggccat	gcaacaaagg	attggcgaga	tagtggctga	540
aatggatggt	cctctccact	gccgcacaga	attctccacc	caggaagagg	agcagcttcg	600
agcccagggc	agcacagact	attctcttag	cagtgggtgac	aagattcgat	tcttcttga	660
gaaaggcggt	tttgatgaga	aaggaaattt	cctggctccct	cggagaaaat	ccatcaacaa	720
aattggccac	gctctgcacg	cccacgaccc	cgtcttcaag	agcatcacac	actccttcaa	780
ggtgcagacc	ttggccagaa	gtctgggcct	ccagatgccc	gtggtggtgc	agagcatgta	840
catctttaag	caacctcact	ttggcgggtga	agtctccctc	catcaggacg	cctccttcct	900
gtacacggag	cccctggggc	gggtgctggg	cgtgtggatc	gcagtggagg	atgccacgct	960
ggagaacggc	tgtctctggt	tcctccctgg	ctccccacac	agtgggtgtg	caagaaggat	1020
ggtccggggc	cctgttggtc	cagcgccctg	taccagcttc	cttgggtcag	agccagcccc	1080

ggataacagc	ctctttgtgc	ccacccagc	gcagagaggg	gccctggctc	tcctccatgg	1140
agaagtggta	cacaagagca	agcagaacct	ctctgaccgc	tcgcgccagg	cctacacttt	1200
ccacctcatg	gaggcctctg	gcaccacctg	gagcccgagg	aactggctcc	agccaacagc	1260
tgaactgccc	tttccccaac	tgtacacctt	aaggtctctg	cagggcagga	gcctctgccc	1320
ctcccgggtg	aagctgtggg	ctgtaaacac	cagtgccttg	ctcagcctcc	tggttgcaac	1380
agggagggtc	tgtctccct	cctgggcttt	cctctgccc	tgtgggcagc	agcctagggt	1440
gggtcagggg	cttcctaag	atcttcacct	ctctgcctcc	ctactgcccc	aacatagcct	1500
tgaggaggct	tctcagccac	caaagggttc	tgcccccttc	tcactctcct	ctcctctcag	1560
atggaactct	ggttattatg	gtgttagtta	tcgaataaaa	acgacttcag	aatgaaaaaa	1620
aaaa						1624

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
21 June 2001 (21.06.2001)

PCT

(10) International Publication Number
WO 01/44448 A3

(51) International Patent Classification⁷: C12N 9/02,
15/10, 15/11, C12Q 1/26, A01K 67/027, A61K 38/44

Valley, CA 94552 (US). LU, Dyung, Aina, M. [US/US];
233 Coy Drive, San Jose, CA 95123 (US).

(21) International Application Number: PCT/US00/33158

(74) Agents: HAMLET-COX, Diana et al.; Incyte Genomics,
Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).

(22) International Filing Date: 7 December 2000 (07.12.2000)

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ,
DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR,
HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR,
LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ,
NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM,
TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/172,367 16 December 1999 (16.12.1999) US

(71) Applicant (*for all designated States except US*): INCYTE
GENOMICS, INC. [US/US]; 3160 Porter Drive, Palo
Alto, CA 94304 (US).

(84) Designated States (*regional*): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European
patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,
IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF,
CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): YUE, Henry
[US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US).
LAL, Preeti [IN/US]; P.O. Box 5142, Santa Clara, CA
95056 (US). TANG, Y., Tom [CN/US]; 4230 Ranwick
Court, San Jose, CA 95118 (US). HILLMAN, Jennifer,
L. [US/US]; 230 Monroe Drive #17, Mountain View, CA
94040 (US). BAUGHN, Mariah, R. [US/US]; 14244
Santiago Road, San Leandro, CA 94577 (US). AZIMZAI,
Yalda [US/US]; 5518 Boulder Canyon Drive, Castro

Published:

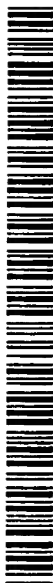
— with international search report

(88) Date of publication of the international search report:
17 January 2002

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: HUMAN OXIDOREDUCTASE PROTEINS

(57) Abstract: The invention provides human oxidoreductase proteins (ORP) and polynucleotides which identify and encode ORP. The invention also provides expression vectors, host-cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with expression of ORP.



WO 01/44448 A3

INTERNATIONAL SEARCH REPORT

In tional Application No

PCT/US 00/33158

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N9/02 C12N15/10 C12N15/11 C12Q1/26 A01K67/027
A61K38/44

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12N A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

STRAND, EPO-Internal, MEDLINE, WPI Data, PAJ, EMBL

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE EMBL [Online] Accession Number AF132944, 27 April 1999 (1999-04-27) LAI C-H ET AL.: "Homo sapiens CGI-10 protein mRNA, complete cds" XP002162312 the whole document	1-15
P,X	--- DATABASE EMBL [Online] Accession Number Q9Y2Z9, 30 May 2000 (2000-05-30) LIN W-C ET AL.: "Putative Ubiquinone biosynthesis monooxygenase COQ6 (EC 1.14.13)" XP002163100 the whole document --- -/-	1,2,10

☒ Further documents are listed in the continuation of box C.☐ Patent family members are listed in annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"Z" document member of the same patent family

Date of the actual completion of the international search

21 March 2001

Date of mailing of the international search report

19.06.01

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Armandola, E

INTERNATIONAL SEARCH REPORT

Int'l Application No

PCT/US 00/33158

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	LAI C-H ET AL.: "Identification of novel human genes evolutionarily conserved in <i>Caenorabditis elegans</i> by comparative proteomics." GENOME RES., vol. 10, no. 5, May 2000 (2000-05), pages 703-713, XP002163099 the whole document ---	1-15
A	BLANCHET M A ET AL.: "Structure and mechanism of cytosolic quinone reductases" BIOCHEM SOC.TRANS., vol. 27, no. 4, August 1999 (1999-08), pages 610-615, XP000984873 -----	

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 00/33158

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

Although claim 18 is directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.
2. ☒ Claims Nos.: 20, 21, 23, 24
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
see FURTHER INFORMATION sheet PCT/ISA/210
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-28 partially

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-28 (partially)

An isolated peptide comprising SEQ. ID. NO: 1, a sequence 90% identical, an active fragment thereof, an immunogenic fragment thereof, an isolated polynucleotide encoding the polypeptide and corresponding to SEQ. ID. NO: 28, cells transformed with the polynucleotide linked to a promoter, a transgenic animal comprising the polynucleotide linked to a promoter, an antibody recognizing the polynucleotide, a method of detecting the polynucleotide by PCR or hybridization, a composition containing the polypeptide, methods for treating a disease by administering the composition, methods for screening agonists and antagonists of the polypeptide, compositions and methods linked to the alteration of the expression of the polypeptide, methods for determining the toxicity of the polypeptide.

2. Claims: 1-28 (partially)

Same as 1. with reference to SEQ. ID. NO: 2 and 29

3. Claims: 1-28 (partially)

Same as 1. with reference to SEQ. ID. NO: 3 and 30

4. Claims: 1-28 (partially)

Same as 1. with reference to SEQ. ID. NO: 4 and 31

5. Claims: 1-28 (partially)

Same as 1. with reference to SEQ. ID. NO: 5 and 32

6. Claims: 1-28 (partially)

Same as 1. with reference to SEQ. ID. NO: 6 and 33

7. Claims: 1-28 (partially)

Same as 1. with reference to SEQ. ID. NO: 7 and 34

8. Claims: 1-28 (partially)

Same as 1. with reference to SEQ. ID. NO: 8 and 35

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

9. Claims: 1-28 (partially)

Same as 1. with reference to SEQ. ID. NO: 9 and 36

10. Claims: 1-28 (partially)

Same as 1. with reference to SEQ. ID. NO: 10 and 37

11. Claims: 1-28 (partially)

Same as 1. with reference to SEQ. ID. NO: 11 and 38

12. Claims: 1-28 (partially)

Same as 1. with reference to SEQ. ID. NO: 12 and 39

13. Claims: 1-28 (partially)

Same as 1. with reference to SEQ. ID. NO: 13 and 40

14. Claims: 1-28 (partially)

Same as 1. with reference to SEQ. ID. NO: 14 and 41

15. Claims: 1-28 (partially)

Same as 1. with reference to SEQ. ID. NO: 15 and 42

16. Claims: 1-28 (partially)

Same as 1. with reference to SEQ. ID. NO: 16 and 43

17. Claims: 1-28 (partially)

Same as 1. with reference to SEQ. ID. NO: 17 and 44

18. Claims: 1-28 (partially)

Same as 1. with reference to SEQ. ID. NO: 18 and 45

19. Claims: 1-28 (partially)

Same as 1. with reference to SEQ. ID. NO: 19 and 46

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

20. Claims: 1-28 (partially)

Same as 1. with reference to SEQ. ID. NO: 20 and 47

21. Claims: 1-28 (partially)

Same as 1. with reference to SEQ. ID. NO: 20 and 48

22. Claims: 1-28 (partially)

Same as 1. with reference to SEQ. ID. NO: 22 and 49

23. Claims: 1-28 (partially)

Same as 1. with reference to SEQ. ID. NO: 23 and 50

24. Claims: 1-28 (partially)

Same as 1. with reference to SEQ. ID. NO: 24 and 51

25. Claims: 1-28 (partially)

Same as 1. with reference to SEQ. ID. NO: 25 and 52

26. Claims: 1-28 (partially)

Same as 1. with reference to SEQ. ID. NO: 26 and 53

27. Claims: 1-28 (partially)

Same as 1. with reference to SEQ. ID. NO: 27 and 54

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 20, 21, 23, 24

Present claims 20, 21, 23 and 24 relate to an extremely large number of possible compounds (agonists/antagonists). In fact, the claims contain so many options that a lack of clarity within the meaning of Article 6 PCT arises to such an extent as to render a meaningful search of the claims impossible.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.